

NBER WORKING PAPER SERIES

A STREAM OF PROSPECTS OR A PROSPECT OF STREAMS:  
ON THE EVALUATION OF INTERTEMPORAL RISKS

James Andreoni  
Paul Feldman  
Charles Sprenger

Working Paper 24075  
<http://www.nber.org/papers/w24075>

NATIONAL BUREAU OF ECONOMIC RESEARCH  
1050 Massachusetts Avenue  
Cambridge, MA 02138  
November 2017

We are grateful to Mark Machina, Yoram Halevy, Levon Barseghyan, P.J. Healy, Marina Agranov, and conference participants at the American Social Science Association Annual Meeting and the Southwest Experimental and Behavioral Economics Conference. Andreoni also thanks the National Science Foundation for financial support. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2017 by James Andreoni, Paul Feldman, and Charles Sprenger. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

A Stream of Prospects or a Prospect of Streams: On the Evaluation of Intertemporal Risks  
James Andreoni, Paul Feldman, and Charles Sprenger  
NBER Working Paper No. 24075  
November 2017  
JEL No. C91,D81,D91

### **ABSTRACT**

Recent debate has identified important gaps in the understanding of intertemporal risks. Critical to closing these gaps is evidence on which dimension of intertemporal risk – the risk or the time – is evaluated first. Though under discounted expected utility this ordering is of no consequence, under discounted non-expected utility models the order of evaluation is critical. We provide experimental tests in which different orderings of evaluation generate different predictions for behavior. We find more support for the notion that the risk dimension is evaluated first.

James Andreoni  
Department of Economics  
University of California, San Diego  
9500 Gilman Drive  
La Jolla, CA 92093-0508  
and NBER  
andreoni@ucsd.edu

Charles Sprenger  
University of California, San Diego  
Rady School of Business  
9500 Gilman Drive  
La Jolla , CA 93093  
c.sprenger@gmail.com

Paul Feldman  
Department of Economics  
University of California, San Diego  
9500 Gilman Drive  
La Jolla , CA 93093  
pfeldman@ucsd.edu

# 1 Introduction

The ability to use the same utility representation for both risk and time preferences has been a defining element of the classic model of Discounted Expected Utility (DEU). This quality, which derives from the assumptions of separability of preferences across time and linearity in probability (the Independence Axiom), has been a key convenience of the model as it implies that desires to smooth consumptions across time and states of nature can be handled symmetrically (see, among others, Hall, 1988; Epstein and Zin, 1991). This feature has also been the focus of a number of important generalizations of DEU (Kreps and Porteus, 1978; Epstein and Zin, 1989; Chew and Epstein, 1989; Epstein and Zin, 1991).

In a recent experimental contribution published in this journal, Andreoni and Sprenger (2012) present subjects with intertemporal decisions where experimental payments are subject to independent risks over time. Their carefully constructed prospects exposed subjects' systematic violations of the DEU assumption that risk preferences and time preferences are interchangeable.<sup>1</sup> Moreover, they go on to claim that their findings are also inconsistent with a number of non-expected utility alternatives, including discounted Cumulative Prospect Theory (CPT) (Kahneman and Tversky, 1979; Tversky and Kahneman, 1992).<sup>2</sup>

---

<sup>1</sup>When allocating a budget across a sooner and a later payment with 100% payment probability, subjects' behavior is substantially different than when each payment will be received with an independent 50% chance, despite the common ratio of payment probabilities. Under DEU, allocating across a sooner and later payment,  $c_1$  and  $c_2$ , each received with corresponding payment probability  $p_1$  and  $p_2$ , subject to intertemporal budget constraint  $c_1 + (1 + r) \cdot c_2 = m$  yields intertemporal Euler equation

$$\frac{u'(c_1)}{D(1, 2) \cdot u'(c_2)} = \left( \frac{p_1}{p_2} \right) \cdot \frac{1}{1 + r},$$

where  $u(\cdot)$  is a stationary instantaneous utility function and  $D(1, 2)$  captures discounting between periods 1 and 2. Consider the two conditions  $(p_1, p_2) = (1, 1)$  and  $(p_1, p_2) = (0.5, 0.5)$ . Given  $p_1/p_2$  is identical, the corresponding Euler equations and allocation behavior should be identical across these conditions. Andreoni and Sprenger (2012) show that individuals allocate more of their budget to the certain  $c_1$  relative to the common ratio counterpart when  $r$  is low and less of their budget to the certain  $c_1$  when  $r$  is high. This cross-over in demand has been reproduced by Miao and Zhong (2015); Cheung (2015) and became the central point of discussion in Epper and Fehr-Duda (2015); Miao and Zhong (2015); Cheung (2015).

<sup>2</sup>A critical component of the CPT model is the assumption of a non-linear probability weighting function,  $\pi(p)$ , rather than the DEU model's assumption of linearity in probabilities. Even with this adjustment, Andreoni and Sprenger (2012) show the Euler equations in the two conditions  $(p_1, p_2) = (1, 1)$  and  $(p_1, p_2) = (0.5, 0.5)$  remain the same because  $\frac{\pi(1)}{\pi(1)} = \frac{\pi(0.5)}{\pi(0.5)} = 1$ . This statement implicitly assumed that CPT is applied atemporally within each period (i.e., a 'risk-first' evaluation of the temporal risk).

After the publication of Andreoni and Sprenger (2012), this journal also published three comments noting potential alternative non-DEU rationalizations of the findings: Miao and Zhong (2015), Cheung (2015) and Epper and Fehr-Duda (2015). Andreoni and Sprenger (2015) replied to Miao and Zhong (2015) and Cheung (2015), acknowledging the general value in their arguments, even though they do not apply to the circumstances of the experiment.<sup>3</sup> However, to fully and fairly examine the very complex and thoughtful hypothesis suggested in the third comment, by Epper and Fehr-Duda (2015), a full experimental test of their unique hypothesis was required. This is the subject of this note.

Epper and Fehr-Duda (2015) argue the Andreoni and Sprenger (2012) findings may be rationalized by CPT if, rather than applying CPT atemporally within each period, as Andreoni and Sprenger (2012) do, one applies the model across periods.<sup>4</sup> This construction of CPT by Epper and Fehr-Duda is novel. Interestingly, the Epper-Fehr-Duda interpretation of intertemporal CPT was not considered or suggested for other prominent intertemporal applications of CPT, such as income targeting (e.g., Camerer, Babcock, Loewenstein and Thaler, 1997; Fehr and Goette, 2007) and myopic loss aversion (e.g., Gneezy and Potters, 1997), whose conceptions of intertemporal CPT are consistent with Andreoni and Sprenger's.<sup>5</sup> As a consequence, Epper

---

<sup>3</sup>Miao and Zhong (2015) and Cheung (2015), propose functional forms in the spirit of models with preferences for the resolution of uncertainty (Kreps and Porteus, 1978; Epstein and Zin, 1989; Chew and Epstein, 1989; Epstein and Zin, 1991) wherein the intertemporal elasticity of substitution and the coefficient of relative risk aversion are governed by different parameters to rationalize the Andreoni and Sprenger (2012) deviations from DEU. As noted in Andreoni and Sprenger (2012) and Andreoni and Sprenger (2015), the majority of models sitting in the above classes of preferences reduce to expected utility when uncertainty is resolved immediately (Kreps and Porteus, 1978; Epstein and Zin, 1989). Given that all uncertainty in Andreoni and Sprenger (2012) was resolved at the end of the experiment, this leaves the researcher in a position of applying the functional forms without the underlying assumptions being satisfied. This point notwithstanding, parametric formulations with assumed differences between risk preferences and time preferences can indeed rationalize the observed findings.

<sup>4</sup>If, instead of applying CPT within period, one applies it across periods, then the two conditions noted in footnote 2 above are qualitatively different. Though in the  $(p_1, p_2) = (1, 1)$  condition only one stream of outcomes can be realized, in the  $(p_1, p_2) = (0.5, 0.5)$  condition, four potential streams exist. The potential streams of outcomes are  $\{(c_1, c_2), (c_1, 0), (0, c_2), (0, 0)\}$  with corresponding probabilities  $\{0.25, 0.25, 0.25, 0.25\}$ . Note, that the ordinal preference ranking of these streams and, hence, the cumulative distribution of streams of outcomes will depend on time preferences and on individual allocations. At low values of  $r$ ,  $(c_1, 0)$  is likely to be ranked above  $(0, c_2)$ , while at high values of  $r$ , the opposite ranking should exist. Given this change in the rankings and corresponding change to the cumulative distribution of streams, one may observe the cross-over in demand observed in the Andreoni and Sprenger (2012) data.

<sup>5</sup> The cab drivers of Camerer et al. (1997) are assumed to view each day's labor market decisions in isolation and the framing effects associated with myopic loss aversion have at their core an implicit assumption that

and Fehr-Duda’s hypothesis represents a potentially important realignment of the literature on intertemporal applications of CPT. CPT’s intertemporal predictions depend on whether temporal risks are perceived as a stream of prospects (‘risk-first’ evaluation, as in Andreoni and Sprenger), or a prospect of streams (‘time-first’ evaluation).<sup>6</sup> It is critical to know which ordering is more consistent with actual behavior.

We present an experiment where, under CPT, behavior is predicted to differ depending on whether evaluation is risk-first or time-first. Our design carries the intuition of the following thought experiment. Consider the decision between two prospects: Prospect S (for safe) is certainty of \$10 paid in a week; and Prospect R (for risky) is a 50% chance of \$20 paid in a week. If both options were composed with a common, independent 90% chance of \$19 paid *in four weeks*, would the initial choices be revised? Under CPT, it depends on whether the time dimension or the risk dimension is evaluated first.

To evaluate the value of choosing the Prospect R under risk-first we apply CPT twice, once in one week over outcomes \$20 and \$0, and again in four weeks over outcomes \$19 and \$0. Because of time separability of the independent gambles, modulo wealth effects, the latter prospect is of no consequence to the earlier choice.

In contrast, under time-first the evaluation of Prospect R applies CPT only once over four streams of outcomes: (\$20,\$19), (\$20,\$0), (\$0,\$19), and (\$0,\$0), where (\$ $x$ , \$ $y$ ) refers to \$ $x$  in one week and \$ $y$  in four weeks. These streams arise with probabilities 0.45, 0.05, 0.45, and 0.05, respectively. Adding the chance of \$19 in four weeks has a discernible effect on behavior. The decisionmaker must consider the 45% chance of a stream better than the prior best outcome of \$20, and the 95% chance of a stream better than the prior worst outcome of \$0. Similarly for evaluating Prospect S the decisionmaker evaluates two streams, (\$10,\$19) and (\$10,\$0), with

---

without explicit bracketing provided by the experimenter, people will evaluate each period’s risk in isolation.

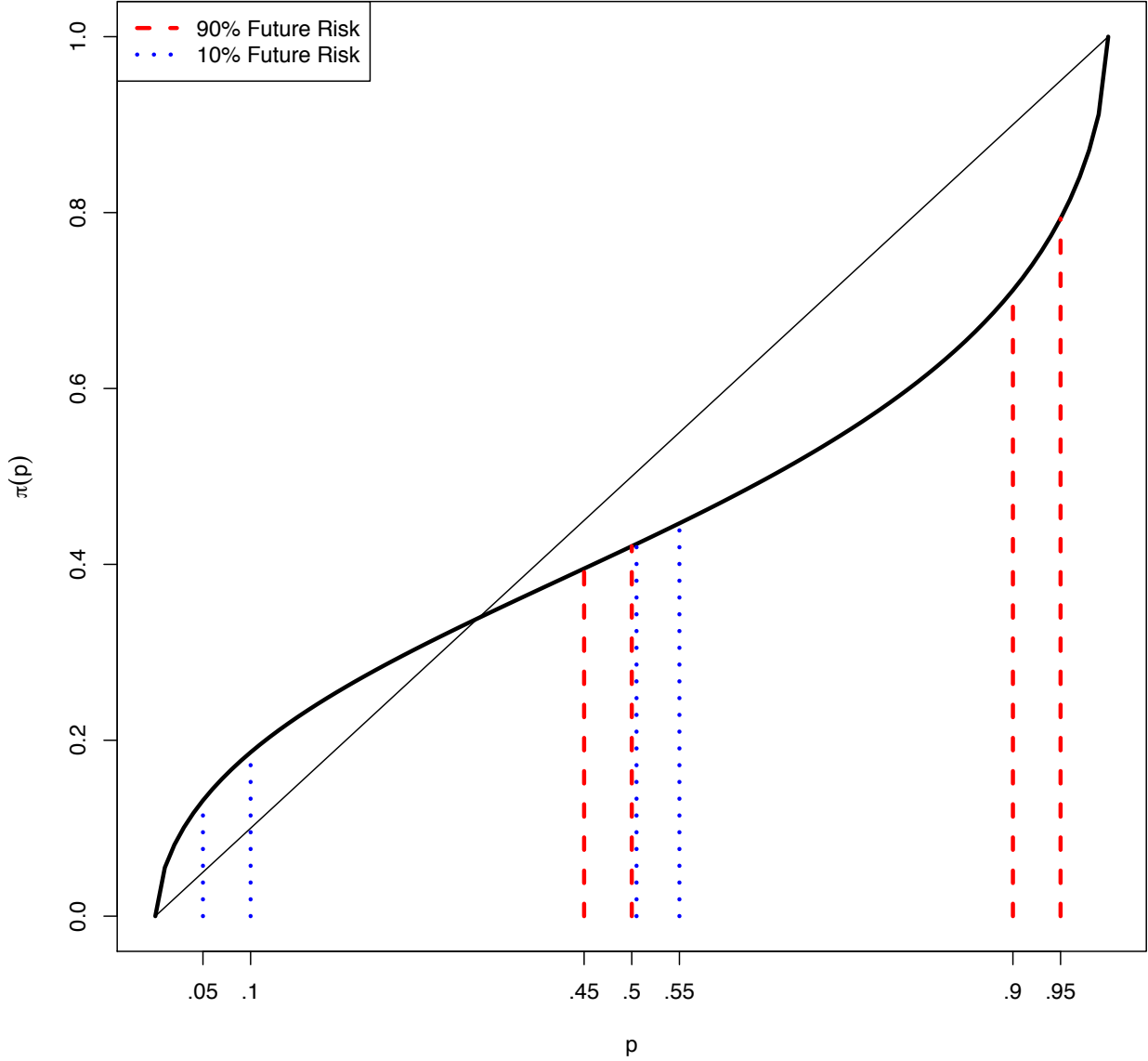
<sup>6</sup>From a normative perspective, both evaluations are consistent with first order stochastic dominance for probability distributions over streams of outcomes. However, when examining a more restrictive notion of consistency, defined as ordinal stochastic dominance (OSD), only the ‘time-first’ formulation satisfies this stronger criterion. Though this may provide some normative appeal for the time-first formulation, it is critical to examine in a simple setting which formulation provides a more compelling account of intertemporal risky choice. A rigorous treatment of OSD appears in section A.2.

a 90% chance of a stream better than the prior best (and worst) outcome of \$10.

Ultimately the new decision between Prospects S and R under the time-first model is arbitrated by a CPT probability weighting function. The central question is whether the difference between 45% and the prior best outcome probability of 50% is larger or smaller than the difference between 90% and 95%. Almost all probability weighting functions proposed in the literature have features of convexity at high probabilities, such that one would expect the difference between 90% and 95% to be larger than between 45% and 50%. Figure 1 provides graphical illustration for the canonical model of probability weighting associated with Tversky and Kahneman (1992). Prospect R benefits from the fact that increasing high cumulative probabilities affects utility more than increasing moderate cumulative probabilities. Hence, a time-first Tversky and Kahneman (1992) decisionmaker who is indifferent to Prospects S and R in isolation would strictly prefer risky Prospect R when adding a common 90% chance of \$19 in four weeks.

Figure 1 highlights the importance of the probabilities associated with common future risks. Most weighting functions are also formulated with concavity at lower probabilities. If the decision between Prospect S and R was taken with a common 10% chance of \$19 in four weeks, the time-first prediction could differ. The calculus then depends upon whether the difference between 5% and 10% is larger or smaller than the difference between 50% and 55%. Here Prospect S benefits from the fact that increasing low cumulative probabilities affects utility more than increasing moderate probabilities. Hence, a time-first Tversky and Kahneman (1992) decisionmaker who is indifferent to Prospects S and R in isolation would strictly prefer safe Prospect S when adding a common 10% chance of \$19 in four weeks. Section 2 and Appendix 1 provide more specific detail and generality.

We treat such kinds of questions as tests of a simple property we call *Common Future Risk Invariance* (CFRI). Whereas the risk-first formulation requires invariance to independent common future risks, the time-first formulation both allows and qualitatively predicts violations



*Notes:* Figure provides standard probability weighting formulation from Tversky and Kahneman (1992),  $\pi(p) = p^\gamma / (p^\gamma + (1 - p)^\gamma)^{1/\gamma}$  and  $\gamma = 0.61$ , with both concave and convex regions. For a time-first decisionmaker indifferent between a safe Prospect S, certainty of \$10 paid in a week, and a risky Prospect R, a 50% chance of \$20 paid a week, the effect of adding a common 90% (10%) chance of \$19 in four weeks is illustrated in red (blue). If  $\pi(0.95) - \pi(0.9) > \pi(0.5) - \pi(0.45)$ , the time-first decisionmaker will strictly prefer the risky Prospect R when adding a common 90% chance of \$19 in four weeks. If  $\pi(0.1) - \pi(0.05) > \pi(0.55) - \pi(0.5)$ , the time-first decisionmaker will strictly prefer the safe Prospect S when adding a common 10% chance of \$19 of four weeks. Both deviations would be predicted by the illustrated weighting function.

Figure 1: Probability Distortions and Time-First Behavior

of this property.<sup>7</sup>

In a sample of 144 undergraduate subjects, we provide within-subject tests of adherence to CFRI, and hence on the order of evaluation in risky temporal choice. Subjects state preferences over two gambles with common, independent future risks. Across tasks, the common future risks are varied both in payment value and in probability.

In addition to testing CFRI, we also confirm that our subjects reproduce the standard findings consistent with CPT probability weighting in an atemporal setting. If indeed the Epper–Fehr–Duda interpretation of CPT is correct, this allows us to make predictions with respect to the CFRI violations; we can forecast their number and sign (greater or lesser risk aversion). From these findings, we predict 70 percent of our subjects should violate CFRI in predictable directions under time-first CPT. Of course, under the risk-first formulation, as well as with DEU, there should be no (systematic or predictable) violations of CFRI.

Violations of CFRI are found in roughly 50 percent of cases in our baseline design. The direction of actual violations is unpredictable, symmetric, and concurring with the prediction of the time-first model at about the rate of chance. Lacking a case for the time-first model, we ask whether the data could be consistent with a formulation of noisy risk-first evaluation. This interpretation is supported by an additional element in our design for which the time-first model predicts effectively no violations of invariance. Nonetheless, violations remain around 50 percent, and remain quite unpredictable relative to the time-first model. As well, the risk-first prediction for behavior generally outperforms the time-first prediction, and particularly when the latter predicts violations of CFRI. Robustness tests assure that the evaluation of risk and time is driving behavior, and not a heuristic of eliminating common decision elements.

Together these data indicate that when risk and time are intertwined, independent risks across periods are evaluated largely in separation. In addition to yielding insights for the interpretation of experimental data, these findings also provide guidance on the application of

---

<sup>7</sup>We view this invariance property as being relevant only for prospects with insignificant wealth effects. If increased probabilistic wealth in the future were to alter the extent of risk taking due to expectations of diminishing marginal utility, even DEU models would yield violations of CFRI.



non-expected utility models in temporal settings. The findings suggest a simple and analytically tractable approach to application when risks are independent. If independent distributions across periods are evaluated separately, then researchers may safely restrict attention only to the periods of interest, assuming that exogenous future risks do not overly influence current choice. This is indeed the assumption that much of the prospect theory literature has implicitly or explicitly invoked to date, as noted above. Such a restriction also disciplines the predictions of CPT through time. As we show, in the time-first application of CPT, changes to common future risks can lead to increased or decreased risk taking, leading many patterns of behavior to be consistent with the model, provided an appropriate formulation of future risks.<sup>8</sup>

Readers who are content with the intuitions provided above may wish to skip section 2 on the detailed theoretical development, going directly to section 3 on experimental procedures. Section 4 provides the results of our experiment and section 5 is a conclusion.

## 2 Theoretical Consideration

Consider a prospect,  $(\mathbf{p}, \mathbf{x})$ , paying outcomes  $\mathbf{x} = \{x_1, x_2, \dots, x_N\}$  with probabilities  $\mathbf{p} = \{p_1, p_2, \dots, p_N\}$ ;  $\sum_{i=1}^N p_i = 1$ . Assume that  $\mathbf{x}$  is a ranked list of gains, relative to reference point  $r \geq 0$ ,  $x_1 \geq x_2 \geq \dots \geq x_N \geq r$ .<sup>9</sup> The CPT utility value for  $(\mathbf{p}, \mathbf{x})$  is

$$U(\mathbf{p}, \mathbf{x}) = \sum_{i=1}^N \left( \pi\left(\sum_{j \leq i} p_j\right) - \pi\left(\sum_{j < i} p_j\right) \right) u(x_i | r),$$

---

<sup>8</sup>The above conclusions are contingent upon wishing to retain the general CPT formulation for risky choice. We provide additional robustness tests implementing common risks within a single period wherein both CPT formulations predict sensitivities. These data highlight general shortcomings in the atemporal predictions of CPT, echoing the recent findings of failures of CPT rank dependence from Bernheim and Sprenger (2017), which may militate for alternative risky choice formulations. Further, even though the data here are supportive of risk-first evaluation, this formulation is clearly rejected by the findings of Andreoni and Sprenger (2012). Resolving these apparent inconsistencies and making positive progress towards more descriptive models of choice in risk over time are important future steps.

<sup>9</sup>Under the Tversky and Kahneman (1992) assumptions, deviations from the reference point — the origin in their formulation — result in both differences in the weighting function and the utility function. Most significant in their formulation is the larger slope in the area below the reference point, when compared to the area above, which induces the kink normally identified as loss aversion. Additional assumptions usually include convexity below the reference point and concavity above the reference point.

where  $u(x_i|r)$  is reference dependent utility, and  $\pi(\cdot)$  is a non-linear probability distortion function acting on cumulative probabilities. When the outcomes  $\mathbf{x}$  are atemporal payments, the application of CPT is straightforward. However, when choices involve both time and risk, the application depends on the order of evaluation of risks through time.<sup>10</sup>

*Risk-First Evaluation:* Consider a CPT decision-maker who faces independent temporal risks and first evaluates each period's distribution of outcomes. Then, in each period,  $t$ , there exists a prospect consisting of a set of outcomes,  $\mathbf{x}_t = \{x_{1,t}, x_{2,t}, \dots, x_{N,t}\}$ , received with probabilities  $\mathbf{p}_t = \{p_{1,t}, p_{2,t}, \dots, p_{N,t}\}$ ;  $\sum_{i=1}^N p_{i,t} = 1$ . When faced with a finite stream of such prospects from  $t = 0$  to  $t = T$ , the decision-maker's utility is

$$\sum_{t=0}^T \left( \sum_{i=1}^N \left( \pi \left( \sum_{j \leq i} p_{j,t} \right) - \pi \left( \sum_{j < i} p_{j,t} \right) \right) \delta^t u(x_{i,t}|r_t) \right),$$

where  $r_t$  is the reference point in each period.

*Time-First Evaluation:* Consider a CPT decision-maker who faces temporal risks and first evaluates each stream of outcomes. In this case, there exists only one prospect  $(\mathbf{p}, \mathbf{X})$ , where the elements of  $\mathbf{X} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_M\}$  are distinct streams of outcomes,  $\{x_0, x_1, \dots, x_T\}$ , which are ordered by their time preference ranking. For example,  $\mathbf{x}_1 = \operatorname{argmax}_{\mathbf{x}} \sum_{t=0}^T \delta^t u(x_t|r_t)$ . The elements of  $\mathbf{p} = \{p_1, p_2, \dots, p_M\}$ ;  $\sum_{i=1}^M p_i = 1$  are the probabilities with which these streams arise. Such a time-first evaluation yields a prospect of streams with utility

$$\sum_{i=1}^M \left( \pi \left( \sum_{j \leq i} p_j \right) - \pi \left( \sum_{j < i} p_j \right) \right) \left( \sum_{t=0}^T \delta^t u(x_{i,t}|r_t) \right).$$

Depending on the order of evaluation, decision-makers should be differentially sensitive to the addition of common future risks. We are specifically interested in the effect of independent common future risks. Consider a prospect,  $(\mathbf{p}_t, \mathbf{x}_t)$ , paying one of  $N$  outcomes in period  $t$

---

<sup>10</sup>In order to focus on the deviations from neoclassical foundations introduced by CPT, our analysis of discounting maintains the benchmark model of exponential discounting (Samuelson, 1937). The conclusions we draw would be maintained under behavioral alternatives to this benchmark such as hyperbolic or quasi-hyperbolic discounting (Laibson, 1997; O'Donoghue and Rabin, 1999). Our design postpones all payments to ensure that immediacy has no additional effect on choice.

alone,  $\mathbf{x}_t = \{x_{1,t}, x_{2,t}, \dots, x_{N,t}\}$ , with probabilities  $\mathbf{p}_t = \{p_{1,t}, p_{2,t}, \dots, p_{N,t}\}$ ;  $\sum_{i=1}^N p_{i,t} = 1$ . Let  $(\mathbf{p}_t, \mathbf{x}_t) + (\mathbf{q}_s, \mathbf{x}_s)$  represent the composition of prospect  $(\mathbf{p}_t, \mathbf{x}_t)$ , with outcomes at time  $t$ , and independent prospect  $(\mathbf{q}_s, \mathbf{x}_s)$ , with outcomes at time  $s > t$ .

For a preference relation on such prospects,  $\succeq$ , we offer the following definition:

**Definition** *Common Future Risk Invariance.* A binary preference relation,  $\succeq$ , exhibits Common Future Risk Invariance (CFRI) if the preference

$$(\mathbf{p}_t, \mathbf{x}_t) \succeq (\mathbf{p}'_t, \mathbf{x}'_t)$$

implies

$$(\mathbf{p}_t, \mathbf{x}_t) + (\mathbf{q}_s, \mathbf{x}_s) \succeq (\mathbf{p}'_t, \mathbf{x}'_t) + (\mathbf{q}_s, \mathbf{x}_s)$$

for all independent future risks  $(\mathbf{q}_s, \mathbf{x}_s)$  in all future periods  $s > t$ .

The risk-first and time-first CPT formulations differentially respect CFRI. The intuition is straightforward. Whereas a risk-first decision-maker considers the stream of prospects  $(\mathbf{p}_t, \mathbf{x}_t)$  and  $(\mathbf{q}_s, \mathbf{x}_s)$  separately through time, the time-first decision-maker evaluates the prospect over streams arrived at by the product of  $(\mathbf{p}_t, \mathbf{x}_t)$  and  $(\mathbf{q}_s, \mathbf{x}_s)$ .

Consider a decisionmaker facing the choices in our introductory thought experiment, who is indifferent between Prospect S, certainty of \$10 paid in a week, and Prospect R, a 50% chance of \$20 paid a week. The CPT representation of this indifference is

$$\delta^1 u(10_1 | r_1) = \delta^1 [\pi(0.5)u(20_1 | r_1) + (1 - \pi(0.5))u(0_1 | r_1)]. \quad (1)$$

Composing these choices with a common 90% chance of receiving \$19 in four weeks leads to our two possible representations. Under the risk-first model, the stream of prospects in the new

choice are again indifferent as

$$\begin{aligned} & \delta^1 u(10_1|r_1) + \delta^4 (\pi(0.9)u(19_4|r_4) + (1 - \pi(0.9))u(0_4|r_4)) = \\ & \delta^1 (\pi(0.5)u(20_1|r_1) + (1 - \pi(0.5))u(0_1|r_1)) + \delta^4 (\pi(0.9)u(19_4|r_4) + (1 - \pi(0.9))u(0_4|r_4)) \end{aligned}$$

reduces to the original equality of (1). In contrast, under the time-first model, the choice requires evaluation of new prospects over ranked streams. For Prospect S, the ranking of streams is  $\mathbf{x}_1 = \{10_1, 19_4\}$ ,  $\mathbf{x}_2 = \{10_1, 0_4\}$ , which arise with probabilities  $\mathbf{p} = \{0.9, 0.1\}$ . The corresponding utility is

$$\pi(0.9) (\delta^1 u(10_1|r_1) + \delta^4 u(19_4|r_4)) + (1 - \pi(0.9)) (\delta^1 u(10_1|r_1) + \delta^4 u(0_4|r_4)) ,$$

or

$$\begin{aligned} & \pi(0.9) \delta^1 u(10_1|r_1) + \\ & \pi(0.9) (\delta^4 u(19_4|r_4)) + (1 - \pi(0.9)) (\delta^4 u(0_4|r_4)) . \end{aligned}$$

For Prospect R, the ranking of streams is  $\mathbf{x}_1 = \{20_1, 19_4\}$ ,  $\mathbf{x}_2 = \{20_1, 0_4\}$ ,  $\mathbf{x}_3 = \{0_1, 19_4\}$ ,  $\mathbf{x}_4 = \{0_1, 0_4\}$ , which arise with probabilities  $\mathbf{p} = \{0.45, 0.05, 0.45, 0.05\}$ . The corresponding utility is

$$\begin{aligned} & \pi(0.45) (\delta^1 u(20_1|r_1) + \delta^4 u(19_4|r_4)) + (\pi(0.5) - \pi(0.45)) (\delta^1 u(20_1|r_1) + \delta^4 u(0_4|r_4)) + \\ & (\pi(0.95) - \pi(0.5)) (\delta^1 u(0_1|r_1) + \delta^4 u(19_4|r_4)) + (1 - \pi(0.95)) (\delta^1 u(0_1|r_1) + \delta^4 u(0_4|r_4)) , \end{aligned}$$

or

$$\begin{aligned} & (\pi(0.5)) (\delta^1 u(20_1|r_1)) + (1 - \pi(0.5)) (\delta^1 u(0_1|r_1)) + \\ & (\pi(0.95) + \pi(0.45) - \pi(0.5)) (\delta^4 u(19_4|r_4)) + (1 - (\pi(0.95) + \pi(0.45) - \pi(0.5))) (\delta^4 u(0_4|r_4)) . \end{aligned}$$

Recalling the indifference from (1), the individual can only remain indifferent if

$$\pi(0.9) (\delta^4 u(19_4|r_4)) + (1 - \pi(0.9)) (\delta^4 u(0_4|r_4)) = (\pi(0.95) + \pi(0.45) - \pi(0.5)) (\delta^4 u(19_4|r_4)) + (1 - (\pi(0.95) + \pi(0.45) - \pi(0.5))) (\delta^4 u(0_4|r_4)),$$

or

$$\pi(0.95) - \pi(0.9) = \pi(0.5) - \pi(0.45).$$

Most probability weighting functions, such as that illustrated in Figure 1 have regions of convexity ranging from moderate to high probabilities. Such a condition would imply the difference between  $\pi(0.95)$  and  $\pi(0.9)$  is larger than that between  $\pi(0.5)$  and  $\pi(0.45)$ . Under such a condition, the individual will prefer Prospect R, growing apparently more risk tolerant under composition with the common future risk.

Importantly, whether common future risk leads to more or less risk tolerance depends upon the relevant probabilities. Consider instead the common future risk being only a 10% chance of receiving \$19 in four weeks. The potential streams and their ranking remains identical for both options, yet the probabilities with which they arise now changes.<sup>11</sup> Now, the decisionmaker

---

<sup>11</sup> For Prospect S, the ranking of streams remains  $\mathbf{x}_1 = \{10_1, 19_4\}$ ,  $\mathbf{x}_2 = \{10_1, 0_4\}$ , which now arise with probabilities  $\mathbf{p} = \{0.1, 0.9\}$ . The corresponding utility is

$$\pi(0.1) (\delta^1 u(10_1|r_1) + \delta^4 u(19_4|r_4)) + (1 - \pi(0.1)) (\delta^1 u(10_1|r_1) + \delta^4 u(0_4|r_4)).$$

For Prospect R, the ranking of streams remains  $\mathbf{x}_1 = \{20_1, 19_4\}$ ,  $\mathbf{x}_2 = \{20_1, 0_4\}$ ,  $\mathbf{x}_3 = \{0_1, 19_4\}$ ,  $\mathbf{x}_4 = \{0_1, 0_4\}$ , which arise with probabilities  $\mathbf{p} = \{0.05, 0.45, 0.05, 0.45\}$ . The corresponding utility is

$$\pi(0.05) (\delta^1 u(20_1|r_1) + \delta^4 u(19_4|r_4)) + (\pi(0.5) - \pi(0.05)) (\delta^1 u(20_1|r_1) + \delta^4 u(0_4|r_4)) + (\pi(0.55) - \pi(0.5)) (\delta^1 u(0_1|r_1) + \delta^4 u(19_4|r_4)) + (1 - \pi(0.55)) (\delta^1 u(0_1|r_1) + \delta^4 u(0_4|r_4)).$$

Indifference requires

$$\begin{aligned} & \delta^1 u(10_1|r_1) + \\ & \pi(0.1) (\delta^4 u(19_4|r_4)) + (1 - \pi(0.1)) (\delta^4 u(0_4|r_4)) = \\ & (\pi(0.5)) (\delta^1 u(20_1|r_1)) + (1 - \pi(0.5)) (\delta^1 u(0_1|r_1)) + \\ & \pi(0.55) - \pi(0.5) + \pi(0.05) (\delta^4 u(19_4|r_4)) + (1 - (\pi(0.55) - \pi(0.5) + \pi(0.05))) (\delta^1 u(0_1|r_1) + \delta^4 u(0_4|r_4)) \end{aligned}$$

will remain indifferent if

$$\pi(0.55) - \pi(0.5) = \pi(0.1) - \pi(0.05).$$

Most parameterizations of CPT with  $S$ -shaped weighting have regions of concavity from low to moderate probabilities. Such a condition would imply the difference between  $\pi(0.55)$  and  $\pi(0.5)$  is smaller than that between  $\pi(0.1)$  and  $\pi(0.05)$ . Under such a condition, the individual will prefer Prospect S, growing apparently more risk averse under composition with the common future risk.

In addition to being sensitive to the probabilities of common future risks, violations of invariance will also be sensitive to the initial probabilities associated with each option. Appendix A.1 provides detailed intuition and simulations evaluating CFRI violations in our experimental environment. One important point for our design is that when Prospect A is initially a 70% chance of receiving \$10, rather than a 100% chance, CFRI violations are predicted to be dramatically less pronounced, to the point of being imperceptible at standard CPT parameter values, and always in the direction of increased risk aversion for the common future risks described above.

Though our study is concerned with assessing the positive content of different formulations for CPT over time, there may also be reasons to evaluate the two formulations on normative grounds.<sup>12</sup> Necessary and sufficient axioms for the time-first evaluation of CPT are provided by

---

By the first indifference between Prospects S and R,

$$\delta^1 u(10_1 | r_1) = \delta^1 [\pi(0.5)u(20_1 | r_1) + (1 - \pi(0.5))u(0_1 | r_1)],$$

and so

$$\begin{aligned} & \pi(0.1) (\delta^4 u(19_4 | r_4)) + (1 - \pi(0.1)) (\delta^4 u(0_4 | r_4)) = \\ & \pi(0.55) - \pi(0.5) + \pi(0.05) (\delta^4 u(19_4 | r_4)) + (1 - (\pi(0.55) - \pi(0.5) + \pi(0.05))) (\delta^1 u(0_1 | r_1) + \delta^4 u(0_4 | r_4)), \end{aligned}$$

$$\pi(0.1) = \pi(0.55) - \pi(0.5) + \pi(0.05).$$

<sup>12</sup>Indeed, CPT was, itself, formulated to address normative concerns related to violations of first order stochastic dominance in original prospect theory. See Quiggin (1982) and Fishburn (1978).

Halevy (2008) and Chew and Epstein (1989). The critical axiom, which distinguishes the risk-first from the time-first formulation, is termed Ordinal Stochastic Dominance (OSD), requiring First Order Stochastic Dominance to be respected for probability distributions over streams ranked according to time preferences over deterministic outcomes. In Appendix A.2, we show that the risk-first formulation violates OSD. We believe the documented violation conveys less about the normative value of the risk-first formulation than about the restrictiveness of OSD, as the examples used are intuitive and behaviorally plausible. Importantly, the time-first formulation, itself, is not without its own normative objections. Ebert and Strack (2015, 2016) show that time-first CPT agents may either bet when assured to eventually lose or never bet when assured an arbitrarily high payoff. Such extreme implications of the time-first formulation are, of course, problematic.

### 3 Experimental Design

Subjects faced two series of five temporal risk tasks. In each task subjects were asked to make a series of decisions between two options. For example, Task 1, reproduced as Figure 2, Panel A, asked subjects to decide between \$10 for sure paid in one week and a changing  $q$ -gamble over \$20 and \$0 paid in one week. The row at which an individual switches from preferring Option A to preferring Option B carries information on the indifference

$$(1; 10_1, 0_1) \sim (q^*; 20_1, 0_1),$$

with  $q^*$  sitting in the interval of probabilities established by the switch point and the subscript 1 indicating the date of payment in one week. In each task, the probability  $q$  generally increased by three percentage points with the objective of establishing a narrow range within which  $q^*$  could be identified. Tasks such as Task 1 serve to provide a benchmark for behavior. Each subsequent task in the series composes these choices with different common future risks. For example, Task 2, reproduced as Figure 2, Panel B, composes the choices of Task 1 with a

common 90% chance of receiving \$19 in four weeks.

TASK 1					
Option A \$10 in a week Chance			<i>or</i>	Option B \$20 in a week Chance	
1)	100 in 100 Chance	<input type="checkbox"/>	<i>or</i>	5 in 100 Chance	<input type="checkbox"/>
2)	100 in 100 Chance	<input type="checkbox"/>	<i>or</i>	10 in 100 Chance	<input type="checkbox"/>
3)	100 in 100 Chance	<input type="checkbox"/>	<i>or</i>	13 in 100 Chance	<input type="checkbox"/>
.					
.					
30)	100 in 100 Chance	<input type="checkbox"/>	<i>or</i>	95 in 100 Chance	<input type="checkbox"/>

Panel A: Task 1

TASK 2							
Option A			<i>or</i>		Option B		
\$10 in a week	\$19 in 4 weeks				\$20 in a week	\$19 in 4 weeks	
Chance 1	Chance 2				Chance 1	Chance 2	
1)	100 in 100 Chance	90 in 100 Chance	<input type="checkbox"/>	<i>or</i>	5 in 100 Chance	90 in 100 Chance	<input type="checkbox"/>
2)	100 in 100 Chance	90 in 100 Chance	<input type="checkbox"/>	<i>or</i>	10 in 100 Chance	90 in 100 Chance	<input type="checkbox"/>
3)	100 in 100 Chance	90 in 100 Chance	<input type="checkbox"/>	<i>or</i>	13 in 100 Chance	90 in 100 Chance	<input type="checkbox"/>
.							
.							
.							
30)	100 in 100 Chance	90 in 100 Chance	<input type="checkbox"/>	<i>or</i>	95 in 100 Chance	90 in 100 Chance	<input type="checkbox"/>

Panel B: Task 2

Figure 2: Sample Tasks

Table 1 outlines two series of temporal risk tasks. In our primary design, Series 1, \$10 paid with certainty in one week is compared to a chance of \$20 in one week. This choice is then composed with a common 90% chance of receiving \$19 in four weeks, 10% chance of receiving \$19 in four weeks, 90% chance of receiving \$9 in four weeks, and 10% chance of receiving \$9 in four weeks. The outcomes are chosen by design such that under weak restrictions on preferences (\$19 in four weeks is worth less than \$20 in one week, and \$9 in four weeks is worth less than \$10



in one week) an explicit ranking of outcomes and, hence, of intertemporal streams of outcomes exists.<sup>13</sup> In the second series of tasks, Option A carries a 70% chance of receiving \$10 in one week,  $(0.7; 10_1, 0_1)$ , rather than certainty of \$10. As noted below, these tasks are predicted to deliver fewer violations of CFRI under the time-first formulation.

The thirty rows of each task provide thirty opportunities for subjects to violate CFRI between any two tasks in a series. With five tasks per series, ten such comparisons of two tasks can be made in each series. We calculate violations of CFRI as the difference in the number of Option A choices between task  $j$  and task  $k$  with  $j < k$ . Negative numbers correspond to violations towards increased risk aversion, positive numbers to violations towards increased risk tolerance.<sup>14</sup>

In addition to these two series of tasks, subjects also completed a series of prospect theory assessment tasks following the design of Tversky and Kahneman (1992) and several robustness tasks which either presented temporal prospects in compounded form and/or removed the temporal dimension entirely. These tasks are described in detail in Appendix A.3. Importantly, the prospect theory assessment tasks allow us to estimate CPT parameters for our specific sample and so to predict the extent of violations of CFRI under the time-first formulation.

One hundred and forty four undergraduate subjects from UC San Diego participated in the study. Six sessions were conducted from 03/31/2016 to 04/28/2016. The tasks were conducted with paper-and-pencil and organized in ‘Task Blocks’ exactly following the series of Table 1. Separate instructions were provided at the beginning of each task block. The full set of instructions can be found in Appendix A.4. In order to provide incentive for truthful revelation

---

<sup>13</sup>An alternative design would be to elicit certainty equivalents for atemporal gambles and compose both with common future risks. Given that the ranking of outcomes is critical to the application of CPT, we felt it prudent to have experimental control of all outcomes. Hence, under relatively weak assumptions on preferences, a known ordering of both outcomes and the streams exists. This design choice comes with costs and benefits. Given the documented increased risk aversion in probability equivalent tasks such as these relative to certainty equivalent tasks of the form generally used to elicit CPT parameters (see, e.g. Hershey, Kunreuther and Schoemaker, 1982; Hershey and Schoemaker, 1985; Schoemaker, 1990; Sprenger, 2015), a mismatch in the levels of risk aversion may exist across experimental environments. Though this may influence the level of risk taking, it is not obvious that this should influence the response to composition with common future risks.

<sup>14</sup>We restrict attention into individuals with unique switch points (see below) such that this measure correctly calculates the exact number of CFRI violations.

Table 1: Experimental Tasks and Mean Response

Series 1 Tasks			Series 2 Tasks		
Task #	Description	Mean $q^*$ (s.e.)	Task #	Description	Mean $q^*$ (s.e.)
1)	(1; 10 <sub>1</sub> , 0 <sub>1</sub> ) <i>vs.</i> (q; 20 <sub>1</sub> , 0 <sub>1</sub> )	0.689 (0.014)	6)	(.70; 10 <sub>1</sub> , 0 <sub>1</sub> ) <i>vs.</i> (q; 20 <sub>1</sub> , 0 <sub>1</sub> )	0.554 (0.012)
2)	(1; 10 <sub>1</sub> , 0 <sub>1</sub> ) <i>vs.</i> (q; 20 <sub>1</sub> , 0 <sub>1</sub> ) + (0.90; 19 <sub>4</sub> , 0 <sub>4</sub> )	0.681 (0.015)	7)	(.70; 10 <sub>1</sub> , 0 <sub>1</sub> ) <i>vs.</i> (q; 20 <sub>1</sub> , 0 <sub>1</sub> ) + (0.90; 19 <sub>4</sub> , 0 <sub>4</sub> )	0.540 (0.014)
3)	(1; 10 <sub>1</sub> , 0 <sub>1</sub> ) <i>vs.</i> (q; 20 <sub>1</sub> , 0 <sub>1</sub> ) + (0.10; 19 <sub>4</sub> , 0 <sub>4</sub> )	0.699 (0.014)	8)	(.70; 10 <sub>1</sub> , 0 <sub>1</sub> ) <i>vs.</i> (q; 20 <sub>1</sub> , 0 <sub>1</sub> ) + (0.10; 19 <sub>4</sub> , 0 <sub>4</sub> )	0.570 (0.012)
4)	(1; 10 <sub>1</sub> , 0 <sub>1</sub> ) <i>vs.</i> (q; 20 <sub>1</sub> , 0 <sub>1</sub> ) + (0.90; 9 <sub>4</sub> , 0 <sub>4</sub> )	0.681 (0.014)	9)	(.70; 10 <sub>1</sub> , 0 <sub>1</sub> ) <i>vs.</i> (q; 20 <sub>1</sub> , 0 <sub>1</sub> ) + (0.90; 9 <sub>4</sub> , 0 <sub>4</sub> )	0.540 (0.013)
5)	(1; 10 <sub>1</sub> , 0 <sub>1</sub> ) <i>vs.</i> (q; 20 <sub>1</sub> , 0 <sub>1</sub> ) + (0.10; 9 <sub>4</sub> , 0 <sub>4</sub> )	0.688 (0.015)	10)	(.70; 10 <sub>1</sub> , 0 <sub>1</sub> ) <i>vs.</i> (q; 20 <sub>1</sub> , 0 <sub>1</sub> ) + (0.10; 9 <sub>4</sub> , 0 <sub>4</sub> )	0.562 (0.012)

*Notes:* Description of experimental tasks for Series 1 and Series 2. Notation  $(p; X_1, 0_1)$  refers to  $p$  chance of receiving \$X in one week or zero otherwise in one week. Notation  $+(p; X_4, 0_4)$  indicates composition with common future risk of  $p$ -chance of receiving \$X in four weeks or zero otherwise in four weeks, applied to both prospects. Mean behavior estimated based on interval regression (Stewart, 1983) of  $q^*$  on indicators for task with standard errors clustered on individual level. Test of null hypothesis for equal behavior across conditions for Series 1 ( $\chi^2(4) = 8.22$ ,  $p = 0.084$ ) and for Series 2 ( $\chi^2(4) = 11.37$ ,  $p = 0.023$ ).

of preferences, subjects were paid according to one, randomly chosen decision in the experiment.<sup>15</sup> On average, subjects earned \$20.40.<sup>16</sup> Comparable to other similar experiments, 30

<sup>15</sup>As our experiment has both risk and time dimensions, the process of payment is somewhat more complicated than standard experiments in either domain separately. When beginning the experiment subjects were notified that they would be receiving their experimental payment by mail and were asked to self-address two envelopes, one for a sooner payment and one for a later payment. In order to eliminate any differential uncertainty or transaction costs across dates, we opted to make all experimental payments at future dates: one week and four weeks from the study date. Further, each subject was given a \$10 participation payment, split between these two dates. Regardless of choice, subjects would always receive \$5 in one week and \$5 in four weeks. These payments, which have grown common in the experimental literature (see, e.g., Andreoni and Sprenger, 2015; Augenblick, Niederle and Sprenger, 2014), are designed to help equalize transaction costs across time. In our context, these payments also help to ensure equality in experimental receipts on the dates when risks are imposed, controlling for potential issues related to differential income effects across dates.

<sup>16</sup>The random-lottery incentive mechanism, which is widely used in experimental economics, does introduce a compound lottery to the decision environment. The theoretical literature on choice under risk, dating to Holt (1986) and Karni and Safra (1987), has suggested that random mechanisms need not be incentive compatible for choices between risky prospects if either the Independence or Reduction of Compound Lotteries axioms are violated. As CPT violates independence, the mechanism may be flawed for the elicitation of CPT behavior. Importantly, however, Starmer and Sugden (1991); Cubitt, Starmer and Sugden (1998) demonstrate that this mechanism can be used even when individuals deviate from expected utility as, in practice, subjects appear to treat decisions effectively in isolation. The data from our prospect theory elicitation tasks may be of critical value here. If standard prospect theory shapes are identified there, then isolation is likely a plausible assumption when evaluating our other tasks as well. We also attempt to encourage isolation by telling subjects to ‘treat each decision as if it could be the one that determines your payments.’

of 144 (20.83 percent) subjects switched multiple times in one or more experimental tasks.<sup>17</sup> An additional 9 subjects violated dominance in one or more prospect theory elicitation tasks, preferring either \$0 with certainty to a given prospect or a prospect to its high outcome with certainty. We exclude such subjects leaving a sample of 105 subjects.<sup>18</sup>

## 4 Results

Table 1 presents average values of  $q^*$  for each task, providing a summary of behavior for our 105 subjects. On average, there appears to be little responsiveness to changes in common future risks in both series of behavior. For our baseline Series 1, the differences in  $q^*$  across tasks are on the order of 1 to 2 percentage points at most. We marginally reject the null hypothesis that all tasks have equal values of  $q^*$  in series 1, ( $\chi^2 = 8.22$ ,  $p = 0.084$ ). In Series 2, though the differences in  $q^*$  are again small in magnitude, we reject equality with greater precision ( $\chi^2 = 11.37$ ,  $p = 0.023$ ). Below (and in Appendix A.1) we demonstrate that violations of CFRI under the time-first model should be substantially less frequent in Series 2 than in Series 1, and should generally be in a different direction than that which is observed.<sup>19</sup>

Small effects on aggregate could mask substantial violations at the individual level. As noted above, each series yields 10 potential comparisons to test CFRI between two tasks. For each subject, in each of these 10 comparisons, we calculate the number of invariance violations by taking the difference in the number of Option A choices between task  $j$  and task  $k$  with  $j < k$ . Again, negative numbers correspond to violations with increased risk aversion in task

---

<sup>17</sup>Around 10 percent of subjects feature multiple switch points in similar price-list experiments (Holt and Laury, 2002; Meier and Sprenger, 2010), and as many as 50 percent in some cases (Jacobson and Petrie, 2009). As such multiple switch points may indicate subject confusion, researchers often exclude such observations or mechanically enforce single switch points (Harrison, Lau, Rutstrom and Williams, 2005).

<sup>18</sup>The results are qualitatively unchanged by the inclusion of such subjects taking their first switch point as the relevant choice and/or ignoring the dominance violations when calculating prospect theory parameters.

<sup>19</sup>Both data series carry some small directional effects with 90% chance of \$19 in the future generating more risk tolerance (lower  $q^*$ ) and 10% chance generating more risk aversion (higher  $q^*$ ). Though the magnitudes are far from those predicted, in Series 1, these effects are in the same direction as predicted by time-first CPT. However, in Series 2, both 90% chance and 10% chance should lead directionally to more risk aversion under time-first CPT, contrary to the data. Given the differences across conditions are actually stronger in Series 2, we hesitate to interpret these directional effects as anything more than wealth effects, perhaps modulated by disappointment or elation with regards to future earnings. See Appendix Figures A1 and A2 for further detail.

$k$  relative to  $j$ , positive numbers correspond to violations with increased risk tolerance, and zero corresponds to no violation in a given comparison. Figure 3 provides histograms of these violations. In our baseline design, Series 1, 43.2 percent of comparisons yield zero violations of CFRI and 62 percent yield 1 or fewer. The violations are symmetric, with 28.5 percent of comparisons yielding violations towards increased risk aversion and 28.3 percent towards increased risk tolerance. Similar patterns are observed in Series 2: 53.1 percent of comparisons have zero CFRI violations and 69.4 percent have 1 or fewer.<sup>20</sup>

Also graphed in Figure 3 are CPT predictions for the frequency of violations under risk-first and time-first formulations. Following Tversky and Kahneman (1992) and Tversky and Fox (1995), we obtain CPT parameters from our prospect theory assessment tasks via a non-linear least squares estimation routine.<sup>21</sup> Our subjects reliably reproduce the standard finding of inverse  $S$ -shaped probability weighting from assessment exercises of this form at parameter values close to other empirical estimates. With aggregate estimates of probability weighting in hand, we construct a prediction on the frequency of CFRI violations that should be observed in each comparison.<sup>22</sup>

---

<sup>20</sup>The violations appear symmetric, with 24.8 percent of comparisons yielding violations towards increased risk aversion and 22.1 percent towards increased risk tolerance.

<sup>21</sup>For these certainty equivalents data, we assume a reference point of  $r = 0$  and specific functional forms for utility and probability weighting such that the indifference condition

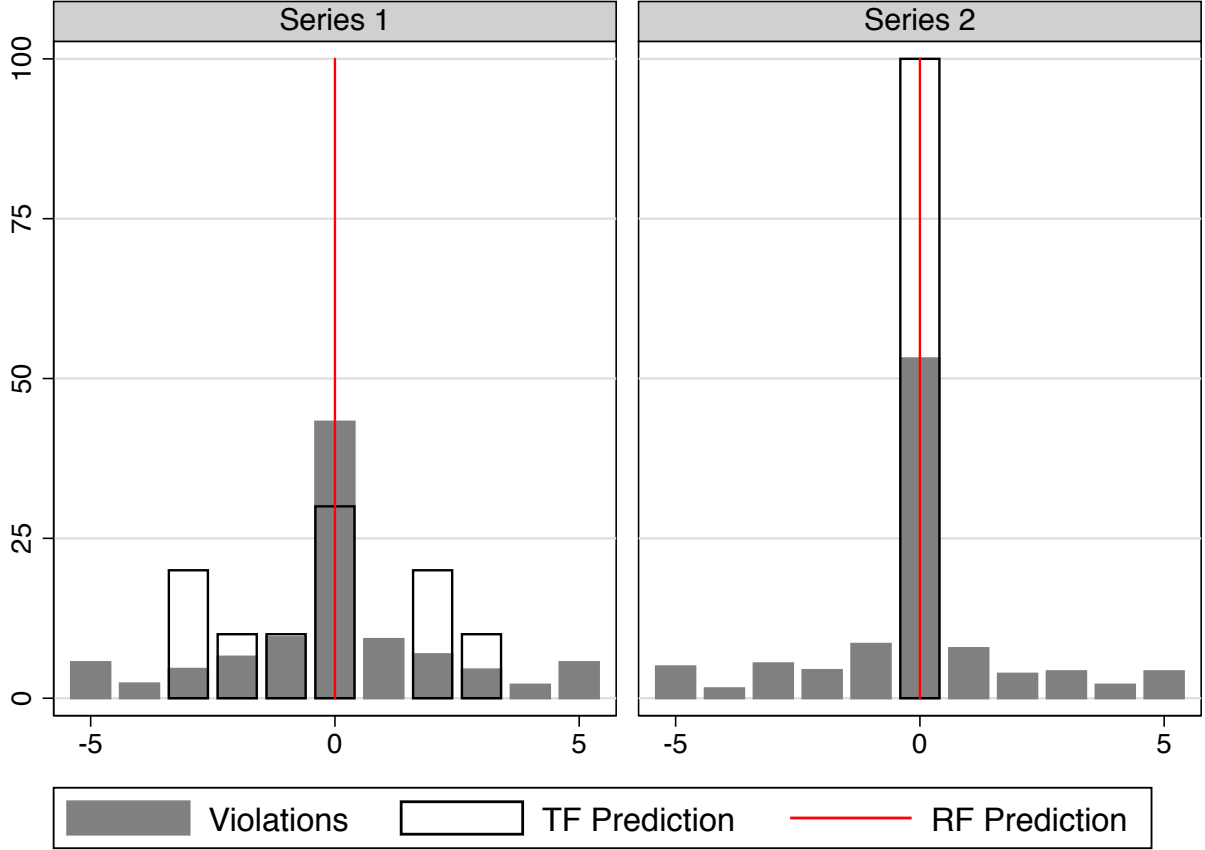
$$\delta^1 u(c|0) = \delta^1 (\pi(p) \cdot u(25|0) + (1 - \pi(p))u(0|0))$$

is met for each observation. We follow the parameterization of Tversky and Kahneman (1992) with power utility above the reference point,  $u(x|0) = (x - 0)^\alpha$ , and the weighting function  $\pi(p) = p^\gamma / (p^\gamma + (1 - p)^\gamma)^{1/\gamma}$ . The parameters  $\hat{\gamma}$  and  $\hat{\alpha}$  are then estimated as the values that minimize the sum of squared residuals of the non-linear regression equation

$$c = [p^\gamma / (p^\gamma + (1 - p)^\gamma)^{1/\gamma} \times 25^\alpha]^{1/\alpha} + \epsilon.$$

where  $c$  is the midpoint of the interval of the certainty equivalent defined by experimental choice. Conducting such analysis on our aggregate data with standard errors clustered on the subject level, we obtain  $\hat{\alpha} = 1.033$  (*s.e.* = 0.031) and  $\hat{\gamma} = 0.696$  (0.017). The benchmark model of expected utility  $\gamma = 1$ , is rejected at all conventional levels, ( $F_{1,104} = 321.32$ ,  $p < 0.01$ ). This compares favorably with other estimates such as Tversky and Kahneman (1992) and Wu and Gonzalez (1996) who estimate  $\hat{\gamma} = 0.61$  and  $\hat{\gamma} = 0.71$ , respectively. Although changes in the parametric specification of the utility function would modify the magnitude of our predicted changes, under most specifications their directions would remain unchanged.

<sup>22</sup>To develop these predictions we apply the probability weighting function previously estimated and allow curvature,  $\alpha$  to be determined by the atemporal choices from Task 1 for Series 1 comparisons and Task 6 for Series 2 comparisons. This allows behavior in Tasks 1 and 6 to serve as the relevant benchmark. These predictions could use the prospect theory assessment tasks' estimate of  $\alpha$ . The implemented procedure allows us to sidestep known differences in risk aversion across probability equivalent and certainty equivalent designs



*Notes:* Histograms of Common Future Risk Invariance violations for Series 1 (Panel A) and Series 2 (Panel B). 10 comparisons per series  $\times$  105 subjects = 1050 observations in each panel. Actual data as grey bars. Time-First Prediction as white bars. Risk-First Prediction of zero violations as red line. All comparisons calculated as difference in number of Option A choices between  $j$  and task  $k$  with  $j < k$ . Negative numbers correspond to increased risk aversion, positive numbers to increased risk tolerance. Actual violations top/bottom coded at  $\pm 5$  for illustration only.

Figure 3: Predicted and Actual CFRI Violations

At the aggregate parameter estimates, violations should be observed in 70 percent of cases in Series 1 for the time-first formulation of CPT. Forty percent of comparisons should yield violations towards increased risk aversion while thirty percent of comparisons should yield violations towards increased risk tolerance. Of course, the risk-first formulation predicts zero violations of CFRI. Non-parametric tests reject equality of the predicted time-first and actual distribution for the frequency of violations ( $z = 2.552$ ,  $p = 0.01$ , Wilcoxon-Mann-Whitney test). The raw correlation between time-first predicted and actual violations is  $\rho = 0.09$  for Series 1. Violations are predicted to be non-existent in Series 2 under the time-first model. Nonetheless actual violations of CFRI remain and the distributions of actual violations are indistinguishable across Series 1 and Series 2 ( $z = 0.757$ ,  $p = 0.45$ , Wilcoxon-Mann-Whitney test). As such, violations of CFRI are symmetric and exist in around 50 percent of cases, whether they are predicted by the time-first model or not.

One first way to assess the relative predictive content of the time-first and risk-first formulations is to compare Mean Squared Error (MSE).<sup>23</sup> In Series 1, the MSE of the time-first formulation is 11.08, while the MSE of the risk-first formulation is 8.16, indicating around a 35% improvement in predictive accuracy when adopting the risk-first formulation. Given that both formulations predict zero violations in Series 2, they share a common MSE value of 9.53.

Prospect theory parameters are estimable at the individual level as well.<sup>24</sup> As such, time-first predictions can be generated for each subject individually. Interestingly, in Series 1 violations of CFRI are predicted to be rather less frequent using individualized measures, only 54.4 percent of cases. Despite these less frequent violations, we again reject equality of distributions between individualized time-first predicted and actual distributions ( $z = 2.305$ ,  $p = 0.02$ , Wilcoxon-Mann-Whitney test). Furthermore, individualized measures do little to improve the correlation between time-first predicted and actual values,  $\rho = 0.09$ , or the relatively poor fit compared to

---

(see, e.g. Hershey et al., 1982; Hershey and Schoemaker, 1985; Schoemaker, 1990; Sprenger, 2015).

<sup>23</sup>Given that the risk-first formulation always predicts zero violation, correlation cannot provide a basis for comparison.

<sup>24</sup>Using prospect theory assessment tasks, we estimate probability weighting,  $\gamma_i$ , for each subject (assuming  $\alpha_i$  equal to 1 for each individual). The distribution of  $\gamma_i$  is provided in Appendix Figure A7. The median value of  $\gamma_i$  is 0.72, echoing the aggregate parameter estimates, and the 5th to 95th percentiles range from 0.41 to 1.02.

the risk-first formulation. The MSE of the time-first formulation reduces only to 10.64 under the individual predictions, again yielding inferior accuracy to the risk-first formulation. Given the heterogeneity in prospect theory parameters, the individual time-first predictions for Series 2 also exhibit some heterogeneity. Violations of CFRI are now predicted in 15.6% of cases under the time-first model. Though the distribution of time-first predictions is not statistically distinguishable from the distribution of actual violations ( $z = 0.86$ ,  $p = 0.39$ , Wilcoxon-Mann-Whitney test), the correlation between predicted and actual violations is again limited,  $\rho = 0.07$ , and the time-first and risk-first models again have effectively identical values of MSE (time-first 9.52, risk-first 9.53).

Though the content of the time-first model's exact predictions are limited, there may be some merit to the model's directional predictions for violations that lead to broadly more or broadly less risk taking. In Table 2, we explore this possibility in more detail. We construct an indicator variable, *Success* ( $=1$ ), which takes the value 1 if a given violation (or lack thereof) is directionally predicted by the time-first formulation and zero otherwise. We regress the measure of predictive success on indicators for whether the time first model predicts a risk averse violation, a risk seeking violation, or no violation at all. Using the aggregate time-first predictions for Series 1, in column (1) we show that the success rate is around 44% when no violation is predicted, and drops to around 35% when the time-first model predicts violations of increased risk aversion or increased risk tolerance. Table 2 also provides comparison of predictive success of the time-first model to random chance. Notably, when the aggregate time-first model predicts a risk averse violation in Series 1 (40% of cases), it actually occurs at a rate that cannot be distinguished from chance ( $F_{1,104} = 0.87$ ,  $p = 0.35$ ).<sup>25</sup> When it predicts a risk seeking violation (30% of cases) it occurs at a rate marginally significantly higher than chance. Overall, the predictive accuracy of the time-first model in Series 1 of 0.352 is around 4 percentage points better than chance, bolstered particularly by the 30%

---

<sup>25</sup>In column (1), the time first formulation predicts 40% risk averse, 30% risk seeking, and 30% no directional violations. Actual directional violations are 28.48%, 28.28%, and 43.24%, respectively. Overall chance success therefor is  $= 100 * (0.4 * 0.2848 + 0.3 * 0.2828 + 0.3 * 0.4324) = 32.85\%$ . Other columns and values are similarly calculated.

Table 2: Predictive Success of Time-First and Risk-First Formulations

<i>Dependent Variable: Success (=1)</i>		Aggregate Predictions		Individual Predictions	
		Series 1 (1)	Series 2 (2)	Series 1 (3)	Series 2 (4)
TF Predicts:					
	Risk Averse Violation	-0.125 (0.061)	- -	-0.132 (0.066)	-0.159 (0.077)
	Risk Tolerant Violation	-0.089 (0.062)	- -	-0.087 (0.066)	-0.279 (0.112)
	No Violation (Constant)	0.441 (0.039)	0.531 (0.034)	0.470 (0.047)	0.544 (0.036)
# Observations		1050	1050	1050	1050
# Clusters		105	105	105	105
<i>Comparison: Time First to Chance Success:</i>					
TF - Chance (TF Predicts R.A. Violation)		0.317-0.285 $F_{1,104} = 0.87$ ( $p = 0.35$ )	- - -	0.337-0.285 $F_{1,104} = 1.73$ ( $p = 0.19$ )	0.385-0.248 $F_{1,104} = 4.84$ ( $p = 0.03$ )
TF - Chance (TF Predicts R.T. Violation)		0.352-0.283 $F_{1,104} = 3.83$ ( $p = 0.05$ )	- - -	0.383-0.283 $F_{1,104} = 6.11$ ( $p = 0.02$ )	0.265-0.221 $F_{1,104} = 0.17$ ( $p = 0.68$ )
TF - Chance (Overall)		0.365-0.328 $F_{1,104} = 4.20$ ( $p = 0.04$ )	- - -	0.409-0.352 $F_{1,104} = 5.21$ ( $p = 0.02$ )	0.515-0.486 $F_{1,104} = 0.88$ ( $p = 0.35$ )
<i>Comparison: Time First to Risk First Success:</i>					
TF - RF (Overall)		0.365-0.432 $F_{1,104} = 14.57$ ( $p < 0.01$ )	- - -	0.409-0.432 $F_{1,104} = 0.91$ ( $p = 0.34$ )	0.515-0.531 $F_{1,104} = 0.27$ ( $p = 0.60$ )
TF - RF (TF Predicts Violation)		0.332-0.429 $F_{1,104} = 11.15$ ( $p < 0.01$ )	- - -	0.357-0.401 $F_{1,104} = 1.85$ ( $p = 0.18$ )	0.360-0.463 $F_{1,104} = 3.08$ ( $p = 0.08$ )

*Notes:* Ordinary least squares regression of *Success (=1)*, successful time-first prediction of CFRI directional violation, in given comparison on constant and predicted directional violation from time-first formulation. Standard errors, clustered, on the individual level in parentheses. Test statistics constructed from linear combinations of regression coefficients tested against chance or risk first predictive success. Column (1): time first formulation predicts 40% risk averse, 30% risk seeking, and 30% no directional violations. Actual directional violations are 28.48%, 28.28%, and 43.24%, respectively. Chance success =  $100 * (0.4 * 0.285 + 0.3 * 0.283 + 0.3 * 0.432) = 32.8\%$ . Other columns and values similarly calculated.



of cases in which no violation is predicted. Quite similar results are observed when using individual predictions in columns (3) and (4), though the predictive accuracy for Series 1 is improved slightly under the individualized measures (largely driven by the increased prediction of non-violations). Interestingly, the individualized measures for Series 2 actually lead to worse predictive accuracy overall (0.531 vs. 0.515), relative to the aggregate Series 2 prediction of zero violations.

Table 2 also provides comparisons between the predictive accuracies of the time-first and risk-first models based on their directional success rates. Overall in Series 1, the risk-first model accuracy exceeds the time-first model by a statistically significant 7 percentage points (around 19 percent), and does so particularly when the time-first model predicts a violation (around 35 percent). Similar conclusions are also drawn when using individualized predictions. When the time-first model deviates from the risk-first prediction of zero CFRI violations, the risk-first model out-predicts it by 4-10 percentage points (11 to 31 percent).

Taken together our results indicate that the risk-first model provides a more coherent account of the data than the time-first model, with noise perhaps generating the CFRI violations that are observed. One plausible explanation for the relative success of the risk-first formulation is that subjects may be eliminating or ignoring common elements across decisions. In two robustness tasks, we repeat the decisions made in Tasks 2 and 3 with all probabilities compounded, removing any common decision elements (see Figure A4, Panel A for an example). Compared to the uncompounded tasks we find both very similar levels of response along with a substantial correlation at the individual level.<sup>26</sup> Though violations of CFRI in these situations are more frequent in the compound frame (89 percent vs. 60 percent), they are again symmetric, and are negatively correlated with the aggregate and individual time-first predictions ( $\rho = -0.165$  and  $\rho = -0.002$ , respectively). The risk-first model again provides enhanced predictive power in terms of MSE for these conditions (32.02 vs. 42.58 and 37.59, for the

---

<sup>26</sup>Task 2's mean  $q^*$  are 0.68 and 0.73 for the uncompounded and compounded tasks, respectively. Task 3's are 0.70 and 0.71. Correlation in number of Option A choices for Task 2 is 0.50 and Task 3 is 0.43 between the compounded and uncompounded tasks.

aggregate and individual predictions, respectively). Appendix Figure A5 provides a histogram of violations for these conditions along with the time first predictions and the corresponding violations in the uncompounded task.<sup>27</sup>

In one further set of robustness tasks, we removed the intertemporal nature of risks. Tasks were identical to Tasks 2-5 and 7-10, except common risks occurred in the same period. In principle, both CPT formulations should predict sensitivity to such common contemporaneous risks. Appendix Figure A6 provides histograms of behavior in these conditions, showing similar behavior when risks are contemporaneous.

## 5 Conclusion

Recognizing that many economic decisions involve both risk and time, we attempt to understand whether the temporal dimension or the risk dimension is evaluated first. Though under the benchmark model of discounted expected utility (DEU) the ordering is of no consequence, under a leading behavioral alternative, Cumulative Prospect Theory (CPT), different orders can lead to qualitative differences in choice behavior. If temporal risks are viewed as a stream of prospects, what we call ‘risk-first’ evaluation, behavior should be uninfluenced by common future risks. In contrast, if such risks are viewed as a prospect of streams, what we call ‘time-first’ evaluation, future risks can have a dramatic influence on behavior, leading to violations of Common Future Risk Invariance (CFRI).

Our experimental examination demonstrates that though the hallmarks of CPT are observed in atemporal choice, violations of invariance predicted by ‘time-first’ evaluation are generally not observed in the data. Violations do occur, but have little predictable sensitivity to common future risks. We interpret these findings as being supportive of noisy ‘risk-first’ evaluation. This interpretation is bolstered by a set of additional tasks and robustness tests demonstrating further lack of predictive accuracy for the time-first model, and that simple heuristics related

---

<sup>27</sup>Interestingly, the distribution of violations across the compounded and uncompounded versions of the tasks are not statistically distinguishable ( $z = -1.319$ ,  $p = 0.19$ , Wilcoxon-Mann-Whitney test).

to the elimination of common attributes are not driving the effects.

These findings have immediate implications for evaluating applications of CPT in temporal settings. Epper and Fehr-Duda (2015) make use of the time-first formulation to rationalize the findings of Andreoni and Sprenger (2012) using CPT. Though such ex-post exercises demonstrate the flexibility of the CPT model, the current results indicate that the behavior documented by Andreoni and Sprenger (2012) is unlikely to be driven by such considerations.

The results may also prove useful for disciplining the application of non-expected utility models through time and for identifying risk preferences in general. If evaluation is ‘time-first’, future exogenous risks can lead to increased risk aversion or increased risk seeking depending on payment probabilities and values. The flexibility of the time-first formulation permits CPT to rationalize many patterns of behavior. Our findings lead naturally to a narrower set of predictions. If future exogenous risks influenced current behavior following time-first logic, the practice of measuring risk preferences would also grow substantially more complex. Researchers would be required to measure not only relevant future risks, but also the heterogeneity thereof, to make statements about the distribution of risk preferences. Furthermore, any apparent instability in measured risk preferences could plausibly be interpreted as changes to future risks. Our findings help to sidestep these natural challenges of the time-first formulation, and help to support existing applications for prospect theory in the field assuming separation through time.

Several extensions exist for our exercise. Most importantly is the design choice of *independent* future risks, uninfluenced by choice. Many natural settings involve current risky decisions that alter the distribution of future risks. Assessing differential formulations of non-DEU models in these settings could yield additional insights.

## References

- Andreoni, James and Charles Sprenger**, “Risk Preferences Are Not Time Preferences,” *American Economic Review*, 2012, *102* (7), 3357–3376.
- and —, “Risk Preferences Are Not Time Preferences: Reply,” *American Economic Review*, 2015, *105* (7), 2287–2293.
- Augenblick, Ned, Muriel Niederle, and Charles Sprenger**, “Working Over Time: Dynamic Inconsistency in Real Effort Tasks,” *Working Paper*, 2014.
- Bernheim, B. Douglas and Charles Sprenger**, “Direct Tests of Cumulative Prospect Theory,” *Working Paper*, 2017.
- Camerer, Colin, Linda Babcock, George Loewenstein, and Richard Thaler**, “Labor Supply of New York City Cabdrivers: One Day at a Time,” *Quarterly Journal of Economics*, 1997, *112* (2), 407–441.
- Cheung, Stephen L.**, “Comment on ‘Risk Preferences Are Not Time Preferences’: On the Elicitation of Time Preference under Conditions of Risk,” *American Economic Review*, 2015, *105* (7), 2242–2260.
- Chew, S. H. and Larry G. Epstein**, “The Structure of Preferences and Attitudes Towards the Timing of the Resolution of Uncertainty,” *International Economic Review*, 1989, *30* (1), 103–117.
- Cubitt, Robin P, Chris Starmer, and Robert Sugden**, “On the validity of the random lottery incentive system,” *Experimental Economics*, 1998, *1* (2), 115–131.
- Ebert, Sebastian and Philipp Strack**, “Until the bitter end: on prospect theory in a dynamic context,” *The American Economic Review*, 2015, *105* (4), 1618–1633.
- and —, “Never, Ever Getting Started: On Prospect Theory Without Commitment,” *Working Paper*, 2016.

- Epper, Thomas and Helga Fehr-Duda**, “Balancing on a Budget Line: Comment on Andreoni and Sprenger (2012)’s ‘Risk Preferences Are Not Time Preferences’,” *American Economic Review*, 2015, 105 (7), 2261–2271.
- Epstein, Larry G. and Stanley E. Zin**, “Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework,” *Econometrica*, 1989, 57 (4), 937–969.
- and —, “Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: An Empirical Analysis,” *Journal of Political Economy*, 1991, 99 (2), 263–286.
- Fehr, Ernst and Lorenz Goette**, “Do Workers Work More if Wages Are High? Evidence from a Randomized Field Experiment,” *The American Economic Review*, 2007, 97 (1), 5.
- Fishburn, Peter C**, “On Handa’s” New theory of cardinal utility” and the maximization of expected return,” *Journal of Political Economy*, 1978, 86 (2, Part 1), 321–324.
- Gneezy, Uri and Jan Potters**, “An Experiment on Risk Taking and Evaluation Periods,” *The Quarterly Journal of Economics*, 1997, 112 (2), 631–645.
- Gonzalez, Richard and George Wu**, “On the Shape of the Probability Weighting Function,” *Cognitive Psychology*, 1999, 38, 129–166.
- Halevy, Yoram**, “Strotz Meets Allais: Diminishing Impatience and the Certainty Effect,” *American Economic Review*, 2008, 98 (3), 1145–1162.
- Hall, Bob**, “Intertemporal Substitution in Consumption,” *Journal of Political Economy*, 1988, 96, 339–357.
- Harrison, Glenn W., Morten I. Lau, Elisabet E. Rutstrom, and Melonie B. Williams**, “Eliciting risk and time preferences using field experiments: Some methodological issues,” in Jeffrey Carpenter, Glenn W. Harrison, and John A. List, eds., *Field experiments*

- in economics*, Vol. Vol. 10 (Research in Experimental Economics), Greenwich and London: JAI Press, 2005.
- Hershey, John C. and Paul J. H. Schoemaker**, “Probability versus Certainty Equivalence Methods in Utility Measurement: Are They Equivalent?,” *Management Science*, 1985, *31* (10), 1213–1231.
- , **Howard C. Kunreuther, and Paul J. H. Schoemaker**, “Sources of Bias in Assessment Procedures for Utility Function,” *Management Science*, 1982, *28* (8), 936–954.
- Holt, Charles A.**, “Preference reversals and the independence axiom,” *The American Economic Review*, 1986, *76* (3), 508–515.
- Holt, Charles A. and Susan K. Laury**, “Risk Aversion and Incentive Effects,” *The American Economic Review*, 2002, *92* (5), 1644–1655.
- Jacobson, Sarah and Ragan Petrie**, “Learning from Mistakes: What Do Inconsistent Choices Over Risk Tell Us?,” *Journal of Risk and Uncertainty*, 2009, *38* (2).
- Kahneman, Daniel and Amos Tversky**, “Prospect Theory: An Analysis of Decision under Risk,” *Econometrica*, 1979, *47* (2), 263–291.
- Karni, Edi and Zvi Safra**, ““ Preference reversal” and the observability of preferences by experimental methods,” *Econometrica: Journal of the Econometric Society*, 1987, pp. 675–685.
- Kreps, David M. and Evan L. Porteus**, “Temporal Resolution of Uncertainty and Dynamic Choice Theory,” *Econometrica*, 1978, *46* (1), 185–200.
- Laibson, David**, “Golden Eggs and Hyperbolic Discounting,” *Quarterly Journal of Economics*, 1997, *112* (2), 443–477.
- Meier, Stephan and Charles Sprenger**, “Present-Biased Preferences and Credit Card Borrowing,” *American Economic Journal - Applied Economics*, 2010, *2* (1), 193–210.

- Miao, Bin and Songfa Zhong**, “Comment on ‘Risk Preference Are Not Time Preferences’: Separating Tisk and Time Preference,” *American Economic Review*, 2015, *105* (7), 2272–2286.
- O’Donoghue, Ted and Matthew Rabin**, “Doing it Now or Later,” *American Economic Review*, 1999, *89* (1), 103–124.
- Prelec, Drazen**, “The Probability Weighting Function,” *Econometrica*, 1998, *66* (3), 497–527.
- Quiggin, John**, “A Theory of Anticipated Utility,” *Journal of Economic Behavior and Organization*, 1982, *3*, 323–343.
- Samuelson, Paul A.**, “A Note on Measurement of Utility,” *The Review of Economic Studies*, 1937, *4* (2), 155–161.
- Schoemaker, Paul J. H.**, “Are Risk-Attitudes Related Across Domains and Response Modes,” *Management Science*, 1990, *36* (12), 1451–1463.
- Sprenger, Charles**, “An Endowment Effect for Risk: Experimental Tests of Stochastic Reference Points,” *Journal of Political Economy*, 2015, *123* (6), 1456–1499.
- Starmer, Chris and Robert Sugden**, “Does the Random-Lottery Incentive System Elicit True Preferences? An Experimental Investigation,” *The American Economic Review*, 1991, *81* (4), 971–978.
- Stewart, Mark B.**, “On Least Squares Estimation when the Dependent Variable is Grouped,” *The Review of Economic Studies*, 1983, *50* (4), 737–753.
- Tversky, Amos and Craig R. Fox**, “Weighing Risk and Uncertainty,” *Psychological Review*, 1995, *102* (2), 269–283.
- **and Daniel Kahneman**, “Advances in Prospect Theory: Cumulative Representation of Uncertainty,” *Journal of Risk and Uncertainty*, 1992, *5* (4), 297–323.

**Wu, George and Richard Gonzalez**, “Curvature of the Probability Weighting Function,”  
*Management Science*, 1996, 42 (12), 1676–1690.



# A Appendix

## A.1 Cumulative Prospect Theory, Violations of Common Future Risk Invariance and Simulated Behavior

*Risk-First Evaluation:* Consider a CPT decision-maker who faces temporal risks and first evaluates each period's distribution of outcomes separately. Then, in each period,  $t$ , there exists a prospect consisting of a set of outcomes,  $\mathbf{x}_t = \{x_{1,t}, x_{2,t}, \dots, x_{N,t}\}$ , received with corresponding probabilities  $\mathbf{p}_t = \{p_{1,t}, p_{2,t}, \dots, p_{N,t}\}$ ;  $\sum_{i=1}^N p_{i,t} = 1$ . When faced with a finite stream of such prospects from  $t = 0$  to  $t = T$ , the decisionmaker's utility is

$$\sum_{t=0}^T \left( \sum_{i,t=1,t}^{N,t} \left( \pi \left( \sum_{j,t \leq i,t} p_{j,t} \right) - \pi \left( \sum_{j,t < i,t} p_{j,t} \right) \right) \delta^t u(x_{i,t} | r_t) \right),$$

where  $r_t$  is reference point in each period.

Importantly, such a decision-maker's atemporal risk taking is unperturbed by composition with common future risks.<sup>28</sup> To illustrate in our context, take an atemporal choice such as that analyzed in Task 1, between \$10 for sure and a  $q\%$  chance of receiving \$20, both to be paid in 1 week's time. Under risk-first evaluation, the indifference condition for  $q^*$  is

$$\delta^1 u(10_1 | r_1) = \delta^1 (\pi(q_1^*) u(20_1 | r_1) + (1 - \pi(q_1^*)) u(0_1 | r_1)).$$

Compose these risks with a common 90% chance of receiving \$19 in four weeks to arrive at the choices analyzed in Task 2, and the indifference condition for  $q^{**}$  is

$$\begin{aligned} \delta^1 u(10_1 | r_1) &+ \delta^4 (\pi(0.9_4) u(19_4 | r_4) + (1 - \pi(0.9_4)) u(0_4 | r_4)) = \\ \delta^1 (\pi(q_1^{**}) u(20_1 | r_1) + (1 - \pi(q_1^{**})) u(0_1 | r_1)) &+ \delta^4 (\pi(0.9_4) u(19_4 | r_4) + (1 - \pi(0.9_4)) u(0_4 | r_4)). \end{aligned}$$

---

<sup>28</sup>Under the risk first evaluation preferences inherit a weak form of intertemporal separability. Essentially, all future common risks do not affect the marginal trade-off rate between the atemporal lotteries.

Noting the elimination of the common 90% of 19 regardless of the shape of utility or probability weighting,

$$q^* = q^{**}.$$

The composition with common future risks will have no effect on choices in our environment because they affect both prospect's valuations equally.

*Time First Evaluation:* Consider a CPT decision-maker who faces the intertemporal risks as above, but first evaluates each stream of outcomes and then evaluates the probability distribution over streams. In this case there exists only one prospect  $(\mathbf{p}, \mathbf{x})$ . The elements of  $\mathbf{x} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N\}$  are now distinct streams of outcomes,  $\{x_0, x_1, \dots, x_T\}$ , which are ordered by their time preference ranking. For example,  $\mathbf{x}_1 = \operatorname{argmax}_{\mathbf{x}} \sum_{t=0}^T \delta^t u(x_t | r_t)$ . The elements of  $\mathbf{p} = \{p_1, p_2, \dots, p_N\}$ ;  $\sum_{i=1}^N p_i = 1$  are the probabilities with which these streams arise. Such a time-first evaluation yields a prospect of streams with utility

$$\sum_{i=1}^N \left( \left( \pi \left( \sum_{j \leq i} p_j \right) - \pi \left( \sum_{j < i} p_j \right) \right) \left( \sum_{i,t=0}^{i,T} \delta^t u(x_{i,t} | r_t) \right) \right),$$

or

$$\sum_{i=1}^N \left( \pi \left( \sum_{j \leq i} p_j \right) - \pi \left( \sum_{j < i} p_j \right) \right) u(\mathbf{x}_i | \mathbf{r}),$$

where  $u(\mathbf{x}_i | \mathbf{r}) = \sum_{i,t=0}^{i,T} \delta^t u(x_{i,t} | r_t)$ .

In contrast to risk-first evaluation, such a time-first decision-maker's atemporal choice behavior should be altered by composition with common future risks.<sup>29</sup> The basic intuition is straightforward. Composition with future risks alters the cumulative distribution of potential streams and so should shift the extent of risk taking. Furthermore, among future risks different payment values and probabilities can have quite different effects. Two future payment values that induce different rankings of streams could lead to quite different levels of risk taking, while two future payment probabilities that induce different cumulative distributions of ranked

---

<sup>29</sup>See appendix 2.

streams could do the same.

Take an atemporal choice such as that analyzed in Task 1, between \$10 for sure and a  $q\%$  chance of receiving \$20, both to be paid in 1 week's time. Under the time first evaluation, the indifference condition for  $q^*$  is

$$\delta^1 u(10_1 | r_1) = \pi(q_1^*) \delta^1 u(20_1 | r_1) + (1 - \pi(q_1^*)) \delta^1 u(0_1 | r_1).$$

Composing these risks with a common 90% chance of receiving \$19 in four weeks to arrive at the choices analyzed in Task 2 now requires the researcher to establish the ranking of streams induced by time preference. Assume  $\delta$ ,  $u(x_t | r_t)$ , and a stream of reference points,  $r_1, r_4$ , such that when considering the  $q\%$  chance of receiving \$20 in one week composed with a 90% chance of \$19 in four weeks, the ranking of streams is  $\mathbf{x}_1 = \{20_1, 19_4\}$ ,  $\mathbf{x}_2 = \{20_1, 0_4\}$ ,  $\mathbf{x}_3 = \{0_1, 19_4\}$ ,  $\mathbf{x}_4 = \{0_1, 0_4\}$ .<sup>30</sup> These streams arise with probabilities  $\mathbf{p} = \{0.9q, 0.1q, 0.9(1 - q), 0.1(1 - q)\}$ . The corresponding utility is

$$\begin{aligned} & \pi(0.9q) (\delta^1 u(20_1 | r_1) + \delta^4 u(19_4 | r_4)) + (\pi(q) - \pi(0.9q)) (\delta^1 u(20_1 | r_1) + \delta^4 u(0_4 | r_4)) + \\ & (\pi(0.9 + 0.1q) - \pi(q)) (\delta^1 u(0_1 | r_1) + \delta^4 u(19_4 | r_4)) + (1 - \pi(0.9 + 0.1q)) (\delta^1 u(0_1 | r_1) + \delta^4 u(0_4 | r_4)). \end{aligned}$$

Similarly, when considering \$10 with certainty in one week composed with a 90% chance of \$19 in four weeks, assume the ranking  $\mathbf{x}_1 = \{10_1, 19_4\}$ ,  $\mathbf{x}_2 = \{10_1, 0_4\}$ , which arise with probabilities  $\mathbf{p} = \{0.9, 0.1\}$ . The corresponding utility is

$$\begin{aligned} & \pi(0.9) (\delta^1 u(10_1 | r_1) + \delta^4 u(19_4 | r_4)) + \\ & (1 - \pi(0.9)) (\delta^1 u(10_1 | r_1) + \delta^4 u(0_4 | r_4)). \end{aligned}$$

---

<sup>30</sup> Assuming the subjects are not facing additional income risks in weeks 2 and 3. For computing our empirical results we assume  $\delta = 1$ ,  $u(x|r) = x^\alpha$  and  $r = 0$ .

The indifference condition for  $q^{**}$  establishes

$$\delta^1 u(10_1|r_1) = \pi(q^{**})\delta^1 u(20_1|r_1) + (\pi(0.9q^{**}) + \pi(0.9 + 0.1q^{**}) - \pi(q^{**}) - \pi(0.9)) (\delta^4 u(19_4|r_4)) + Z,$$

where  $Z$  is a collection of utility terms where the decision-maker receives zero in a given period.<sup>31</sup>

Letting  $u(0|r_t) = 0$ , and recalling the indifference condition from above, this implies,

$$(\pi(q_1^*) - \pi(q^{**})) \delta^1 u(20_1|r_1) = ([\pi(0.9 + 0.1q^{**}) - \pi(0.9)] - [\pi(q^{**}) - \pi(0.9q^{**})]) (\delta^4 u(19_4|r_4)).$$

Hence,  $q^*$  and  $q^{**}$  will only generally be equal if  $\pi(\cdot)$  is linear. Furthermore, the difference between  $q^*$  and  $q^{**}$  will depend on whether adding  $0.1q^{**}$  has a larger impact on probability distortion when probabilities are in the neighborhood of 0.9 or are in the neighborhood of  $q^{**}$ . This depends on both the shape of probability weighting and on the curvature of the utility function. Parametric formulations for  $S$ -shaped probability weighting generally feature a convex region from around probability 0.3 to 1 (Tversky and Kahneman, 1992). If  $q^{**}$  sits in this region but lies below 0.9, the addition of  $0.1q^{**}$  will have a larger influence around probability 0.9, and hence the difference between  $q^*$  and  $q^{**}$  will be positive. Stated differently, the decision-maker will likely grow more risk tolerant when the atemporal risks of Task 1 are composed with common future risks of a 90% chance of \$19 in four weeks in Task 2.

Importantly, given that probability distortions are often posited to have both concave and convex regions (Kahneman and Tversky, 1979), the prediction that individuals will grow more risk tolerant with common future risks is not universal. Consider, instead of a 90% chance of \$19 in 4 weeks, a 10% chance of \$19 in 4 weeks as in Task 3. The ranking of outcomes does not change, but now when considering a  $q\%$  chance of receiving \$20 in one week composed with a 10% chance of receiving \$19 in four weeks,  $\mathbf{p} = \{0.1q, 0.9q, 0.1(1 - q), 0.9(1 - q)\}$ .<sup>32</sup> Let  $q^{***}$  reflect indifference in Task 3. Following the same logic as above one arrives at the following

---

<sup>31</sup> $Z = (\pi(.9) + 1 - \pi(.9q^{**}) - \pi(q + .9(1 - q)))\delta^1 u(0|r) + (1 - \pi(p_2))\delta^4 u(0|r).$

<sup>32</sup>That is, the ranking of streams is  $\mathbf{x}_1 = \{20_1, 19_4\}$ ,  $\mathbf{x}_2 = \{20_1, 0_4\}$ ,  $\mathbf{x}_3 = \{0_1, 19_4\}$ ,  $\mathbf{x}_4 = \{0_1, 0_4\}$ . The

difference between risk taking in Tasks 1 and 3:

$$(\pi(q_1^*) - \pi(q^{***})) \delta^1 u(20_1|r_1) = ([\pi(0.1 + 0.9q^{***}) - \pi(0.1)] - [\pi(q^{***}) - \pi(0.1q^{***})]) (\delta^4 u(19_4|r_4)) .$$

The question now is whether the addition of  $0.9q^{***}$  has a larger effect at probability 0.1 or at probability  $0.1q^{***}$ . Given that 0.1 and  $0.1q^{***}$  are likely to be in the region of concavity then, provided  $\pi(\cdot)$  does not become too sharply convex, the difference is likely to be negative. Stated differently, the decisionmaker will likely grow more risk averse when the atemporal risks of Task 1 are composed with common future risks of a 10% chance of \$19 in four weeks in Task 3.

Developing further intuition about the direction and magnitude of these effects is challenging as predictions are dependent upon not only the exact shape of probability weighting, but also on the combined effects of the utility function and probability distortions. In order to gain further insights on the effects of common future risks, we conduct simulations using the functional forms for probability weighting and utility proposed in the original formulation of CPT (Tversky and Kahneman, 1992)<sup>33</sup>, which posited probability weighting function,  $\pi(p) = p^\gamma / (p^\gamma + (1-p)^\gamma)^{1/\gamma}$ , a reference point of  $r = 0$ , and utility function  $u(x) = x^\alpha$  for  $x > r = 0$ . The parameters identified by Tversky and Kahneman (1992) were  $\gamma = 0.61$  and  $\alpha = 0.88$ . Under these parameters we simulate behavior in Series 1 and Series 2 of our study. The corresponding predictions for behavior are provided in Figure A1, Panels A and B. In each panel we present the atemporal choice as the dashed black line. Given that the risk first evaluation predicts no sensitivity to common future risks, the dashed black line also corresponds to the risk-first prediction.

---

corresponding utility is

$$\begin{aligned} & \pi(0.1q) \quad (\delta^1 u(20_1|r_1) + \delta^4 u(19_4|r_4)) \quad + \\ & (\pi(q) - \pi(0.1q)) \quad (\delta^1 u(20_1|r_1) + \delta^4 u(0_4|r_4)) \quad + \\ & (\pi(0.1 + 0.9q) - \pi(q)) \quad (\delta^1 u(0_1|r_1) + \delta^4 u(19_4|r_4)) \quad + \\ & (1 - \pi(0.1 + 0.9q)) \quad (\delta^1 u(0_1|r_1) + \delta^4 u(0_4|r_4)) . \end{aligned}$$

<sup>33</sup>Tversky and Fox (1995) and Gonzalez and Wu (1999) employ a similar two parameter  $\pi(p)$  function. See Prelec (1998) for alternative S-shaped specifications.

Panel A echoes the intuition presented above. When common future risks occur with high probability decisionmakers grow more risk tolerant, while when they occur with low probability decisionmakers grow more risk averse. The magnitudes can be dramatic. A CPT decisionmaker with  $\gamma = 0.61$  and  $\alpha = 0.88$  should be indifferent between \$10 for sure in a week and a 71.4% chance of \$20 in a week. When composed with a 90% chance of receiving \$19 in four weeks, this probability falls to 60.2%. When composed with a 10% chance of receiving \$19 in four weeks, this probability rises to 72.5%. The existence of future risks can alter risk taking by as much as sixteen percent, and whether future risks are high or low probability can alter risk taking by as much as twenty percent,  $(72.5 - 60.2) / 60.2 = 0.204$

Interestingly, the distinction between high and low probability seems to matter qualitatively more than the distinction between high and low payment values. Given a probability of future payment, both high and low values lead directionally to similar changes in risk taking.

It should be noted that the links between future payment probabilities and risk taking are not universal. In addition to depending on the shape of probability weighting and utility, the predictions also depend on the initial atemporal prospects. As demonstrated in Figure A1, Panel B corresponding to Series 2 of the study, if instead of certainty of \$10, decisionmakers chooses between a 70% chance of \$10 and a  $q\%$  chance of \$20 in one week, common future risks increase risk tolerance regardless of their probability. In Series 2, the effect of common of future risks is smaller in magnitude. Most comparisons between two tasks yield deviations smaller than our experimental granularity of three percentage points. Of additional note is that the relationship between risk taking and future payment values can be non-monotonic. The combined effects of utility function curvature and probability weighting can lead to situations where higher payment values move probability equivalents into regions of greater convexity, generating the potential for such effects.

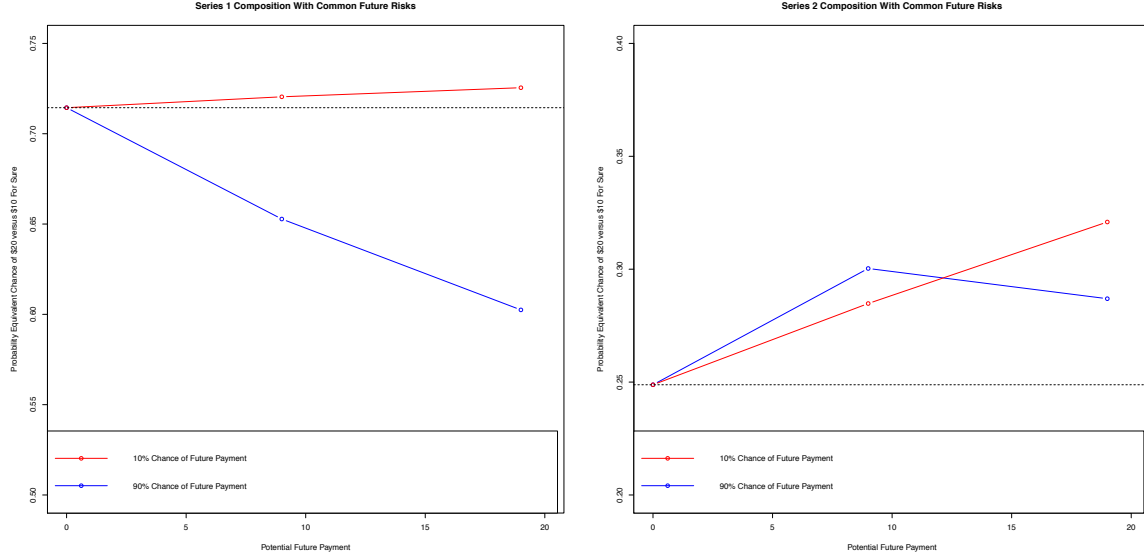


Figure A1: Theoretical Predictions

Whether temporal risks are evaluated in a risk-first or a time-first mode can lead to both qualitative and quantitative differences in predicted behavior. Our experimental design is predicated on investigating the extent to which behavior adheres to each of these models at both the aggregate and individual level. In Figure A2 we provide the time-first predictions and behavior for our sample at the aggregate estimate probability weighting parameter of  $\gamma = 0.697$ . To develop predictions we estimate  $\alpha$  separately for Task 1 and Task 6 following

$$\pi(q) = \frac{10^\alpha}{20^\alpha} + \epsilon,$$

for Task 1 or

$$\pi(q) = \pi(0.7) \frac{10^\alpha}{20^\alpha} + \epsilon,$$

for Task 6. The values  $\pi(q)$  and  $\pi(0.7)$  are the distorted probability values at  $\gamma = 0.697$  for each subject. The objective of this exercise, to let Tasks 1 and 6 serve as benchmarks for their relevant series, seems to be met. At the estimated parameter values, Task 1 and Task 6 predictions are within 0.015 of the true values.

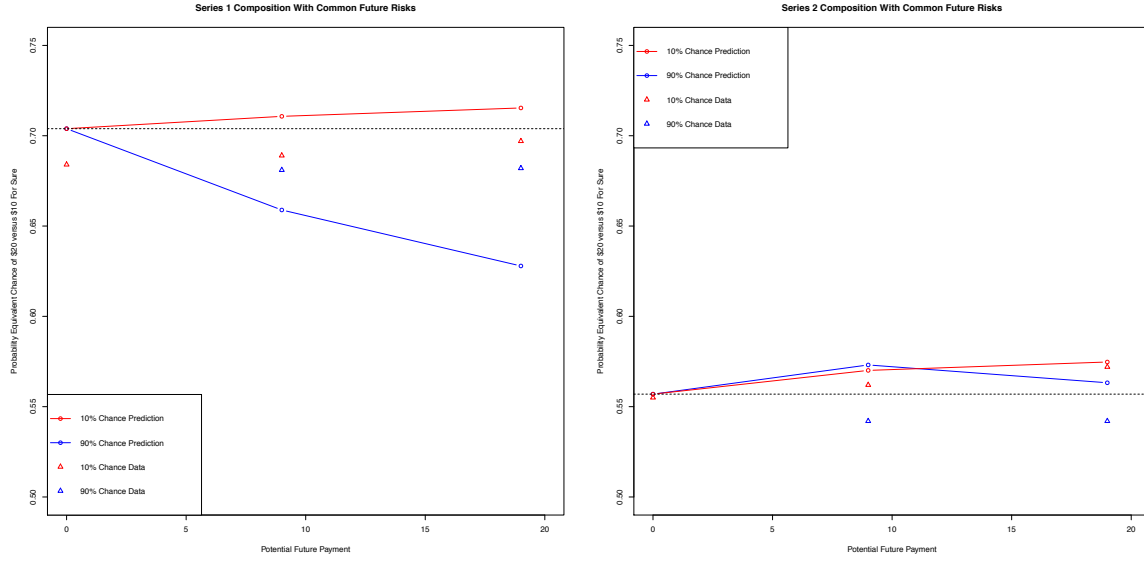


Figure A2: Predictions and Behavior

Our time-first behavioral predictions at the aggregate parameter estimate of  $\gamma = 0.697$  are quite similar to those noted above. In Series 1, sharp differences in behavior are predicted with changes to common future risk, while in Series 2 the sensitivity of behavior is more muted. Indeed, in Series 2, none of the predicted differences are large enough to be identified by our design's granularity of three percentage points. Figure A2 also provides the mean behavior presented in Table 1 for each series. Behavior in Series 1 appears largely unresponsive to future risks. As noted in the text, no statistical differences across tasks in Series 1 are observed. However, the directional predictions do seem consistent, with more risk aversion under the 10% future risks than 90% future risks. Differences across tasks in Series 2, though predicted to be muted, are actually statistically distinguishable. Importantly, the direction of actual differences is inconsistent with the direction of time-first predictions. Instead of growing more risk averse with common future risks, subjects grow somewhat less so for the case of 90% future risks.



## A.2 Ordinal Stochastic Dominance and Normative Implications

This paper is mainly concerned with positive tests of different versions of models of temporal CPT. However, it is worth stressing that CPT was originally developed to address normative concerns. In its original formulation PT lead to violations of first order stochastic dominance<sup>34</sup>. Following this tradition, we would like to stress, that from a normative perspective, time-first can have more desirable normative properties. Since some nuance is involved, the following axiom and proofs clarify the sense in which time-first is more normatively appealing than risk-first.

The necessary and sufficient axioms necessary for a representation of time-first are provided by Halevy (2008) and Chew and Epstein (1989). The critical axiom, for the purpose of this section, is ordinal (stochastic) dominance axiom. This axiom, formally stated below, ensures that first order stochastic dominance is extended and preserved for a given preference ranking over all deterministic time-dated outcomes. As the following example will prove, risk-first violates this axiom.

*Ordinal (Stochastic) Dominance Axiom (OSD):* Let  $\mathcal{C}_+$  be the space of bounded non-negative consumption vectors. For all  $p', p$ , probability distributions over consumption vectors, if  $p'\{c \in \mathcal{C}_+ : c \succeq \bar{c}\} \geq p\{c \in \mathcal{C}_+ : c \succeq \bar{c}\}$  for all  $\bar{c} \in \mathcal{C}_+$  then  $p' \succeq p$ .

*Proof (that risk-first violates Ordinal Dominance)* Let  $L = \{(1; x_1, 0_1), (1; 0_2), (1; 0_3), \dots\}$  and  $L' = \{(1/n; x_{1,1}, 0_1), \dots, (1/n; x_{n,n}, 0_n)\}$  such that  $u(x_{t,t}) = \delta^{-t}u(x - \epsilon)$  and all positive prizes in  $L'$  are perfectly correlated. That is,  $L$  delivers  $x$  units of the consumption good today for sure and  $L'$  will deliver  $x - \epsilon$  worth, in terms of today's present discounted value, of the consumption good during some time period, either today or in the future before some time  $n$ . Clearly,  $L$  OSD  $L'$ , but for a risk-first agent  $U(L) = u(x) < \sum_t \pi(\frac{1}{n})u(x - \epsilon) = U(L')$  for some  $n$  as long as  $\pi$  is concave for low enough values

We do not view this example as particularly strained: why isn't it reasonable to pay an  $\epsilon$  for the surprise of timing? Similar examples going the other direction do not

---

<sup>34</sup>See Quiggin (1982) and Fishburn (1978).

seem obviously strained either. Consider  $L = \{(1; x_1, 0_1), (1; 0_2), (1; 0_3), \dots\}$  and  $L' = \{(1/2; x_1, 0_1), (1/2; x_2, 0_2) + (1; 0_3), \dots\}$ , again with perfect correlation for the outcomes of  $L'$  and  $u(x_1 + \epsilon) = \delta u(x_2)$ .  $L'$  OSD  $L$ , but for the risk-first agent if  $\pi(0.5)$  is sufficiently below 0.5,  $U(L') = \pi(0.5)u(x_1) + \pi(0.5)u(x_1 + \epsilon) < u(x_1) = u(L)$ . Again, it does not seem absurd to pay an  $\epsilon$  to be assured the payment will occur in the first period.<sup>35</sup>

### A.3 Additional Tasks and Behavior

#### A.3.1 Prospect Theory Assessment Tasks

TASK 26					
Option A			<i>or</i>	Option B	
50 in 100 Chance				100 in 100 Chance	
1)	\$25 in a week	<input type="checkbox"/>	<i>or</i>	\$25.00 in a week	<input type="checkbox"/>
2)	\$25 in a week	<input type="checkbox"/>	<i>or</i>	\$24.00 in a week	<input type="checkbox"/>
3)	\$25 in a week	<input type="checkbox"/>	<i>or</i>	\$23.00 in a week	<input type="checkbox"/>
.					
.					
26)	\$25 in a week	<input type="checkbox"/>	<i>or</i>	\$0.00 in a week	<input type="checkbox"/>

Figure A3: CPT Task

In order to provide a prediction for how much the composition of future risks should alter behavior depending on mode of evaluation, we also conduct additional tasks, standard in the CPT literature (Tversky and Kahneman (1992)), for evaluating the extent of probability weighting. Each subject faced 7 tasks eliciting their certainty equivalent to be paid in one week for a  $p\%$  chance of receiving \$25 in one week. Following Tversky and Kahneman (1992), the seven tasks corresponded to  $p \in \{0.05, 0.10, 0.25, 0.50, 0.75, 0.90, 0.95\}$ . An example task is provided as Figure A3 and the seven prospect theory assessment tasks are enumerated in Table 1, Panel

<sup>35</sup>Uncorrelated examples can also be constructed, but are dependent on the shape of the probability weighting function.

### A.3.2 Compounded Tasks and Common Contemporaneous Risk

In addition to the above tasks assessing prospect theory and the mode of evaluation for temporal risks, our design includes two types of robustness tests. First, our tasks are designed in price list format where common future risks are presented as an additional column. Readers may be concerned that heuristics such as the elimination of common attributes might lead common future risks to be ignored. In our first set of robustness tasks, we re-conduct Tasks 2, 3, 11 and 12 with all risks compounded. In such tasks, one cannot eliminate common future risks by simply ignoring a column of the task.<sup>37</sup> Appendix Figure A4, Panel A provides an example. Second, our tasks compose atemporal risky choices with common future risks. A natural question is whether common risks applied in the same period deliver qualitatively different results. Under CPT, regardless of the order of evaluation, the composition with common risks within a period should alter behavior. Our second set of robustness tasks re-conduct tasks 2-5 and 7-10 with the common risks imposed within the same period in one week (as opposed to in four weeks). Appendix Figure A4, Panel B provides an example.<sup>38</sup>

---

<sup>36</sup>We opted to elicit the extent of atemporal non-expected utility behavior at only one point in time, in one week, rather than both in one week and in four weeks. Naturally, if there exist non-stationarities in probability distortions one could imagine that future risks would be treated differently than sooner risks. In the time first evaluation this is difficult to think through ... what function do you use for the cumulatives. Importantly, however, we can examine the relationship between our findings of probability distortions at one week and those at other time scales to assess the plausibility of such lack of stationarity.

<sup>37</sup>It should be noted that subjects were more likely to provide multiple switch points for this task, which could be interpreted as a measure of confusion.

<sup>38</sup>For completeness we also did the common atemporal risks with all the compounding, but we don't talk about that.

## TASK 21

	Option A				Option B				
	\$10 in a week and \$19 in 4 weeks	\$10 in a week and \$0 in 4 weeks		<i>or</i>	\$20 in a week and \$19 in 4 weeks	\$20 in a week and \$0 in 4 weeks	\$0 in a week and \$19 in 4 weeks	\$0 in a week and \$0 in 4 weeks	
1)	90 in 100 Chance	10 in 100 Chance	<input type="checkbox"/>	<i>or</i>	4.5 in 100 Chance	.5 in 100 Chance	85.5 in 100 Chance	9.5 in 100 Chance	<input type="checkbox"/>
2)	90 in 100 Chance	10 in 100 Chance	<input type="checkbox"/>	<i>or</i>	9 in 100 Chance	1 in 100 Chance	81 in 100 Chance	9 in 100 Chance	<input type="checkbox"/>
3)	90 in 100 Chance	10 in 100 Chance	<input type="checkbox"/>	<i>or</i>	11.7 in 100 Chance	1.3 in 100 Chance	78.3 in 100 Chance	8.7 in 100 Chance	<input type="checkbox"/>
.									
.									
30)	90 in 100 Chance	10 in 100 Chance	<input type="checkbox"/>	<i>or</i>	85.5 in 100 Chance	9.5 in 100 Chance	4.5 in 100 Chance	.5 in 100 Chance	<input type="checkbox"/>

(a) Task 21

## TASK 11

	Option A				Option B		
	\$10 in a week Chance 1	\$19 in a week Chance 2		<i>or</i>	\$20 in a week Chance 1	\$19 in a week Chance 2	
1)	100 in 100 Chance	90 in 100 Chance	<input type="checkbox"/>	<i>or</i>	5 in 100 Chance	90 in 100 Chance	<input type="checkbox"/>
2)	100 in 100 Chance	90 in 100 Chance	<input type="checkbox"/>	<i>or</i>	10 in 100 Chance	90 in 100 Chance	<input type="checkbox"/>
3)	100 in 100 Chance	90 in 100 Chance	<input type="checkbox"/>	<i>or</i>	13 in 100 Chance	90 in 100 Chance	<input type="checkbox"/>
.							
.							
30)	100 in 100 Chance	90 in 100 Chance	<input type="checkbox"/>	<i>or</i>	95 in 100 Chance	90 in 100 Chance	<input type="checkbox"/>

(b) Task 11

Figure A4: Sample Robustness Tasks

### A.3.3 Results

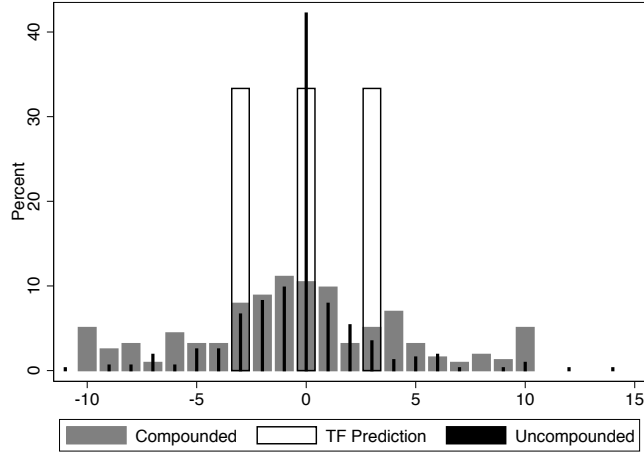


Figure A5: Predicted and Actual CFRI Violations (Compounded)

*Notes:* Histogram of Common Future Risk Invariance violations for Task 1, Task 2 (compounded), and Task 3 (compounded). Three comparisons  $\times$  105 subjects = 315 observations in each panel. Actual data as grey bars. Time-First Prediction as white bars. All comparisons calculated difference in number of Option A choices between task  $j$  and task  $k$  with aggregate prediction for  $j < k$ . Negative numbers correspond to increased risk aversion, positive numbers to increased risk tolerance.

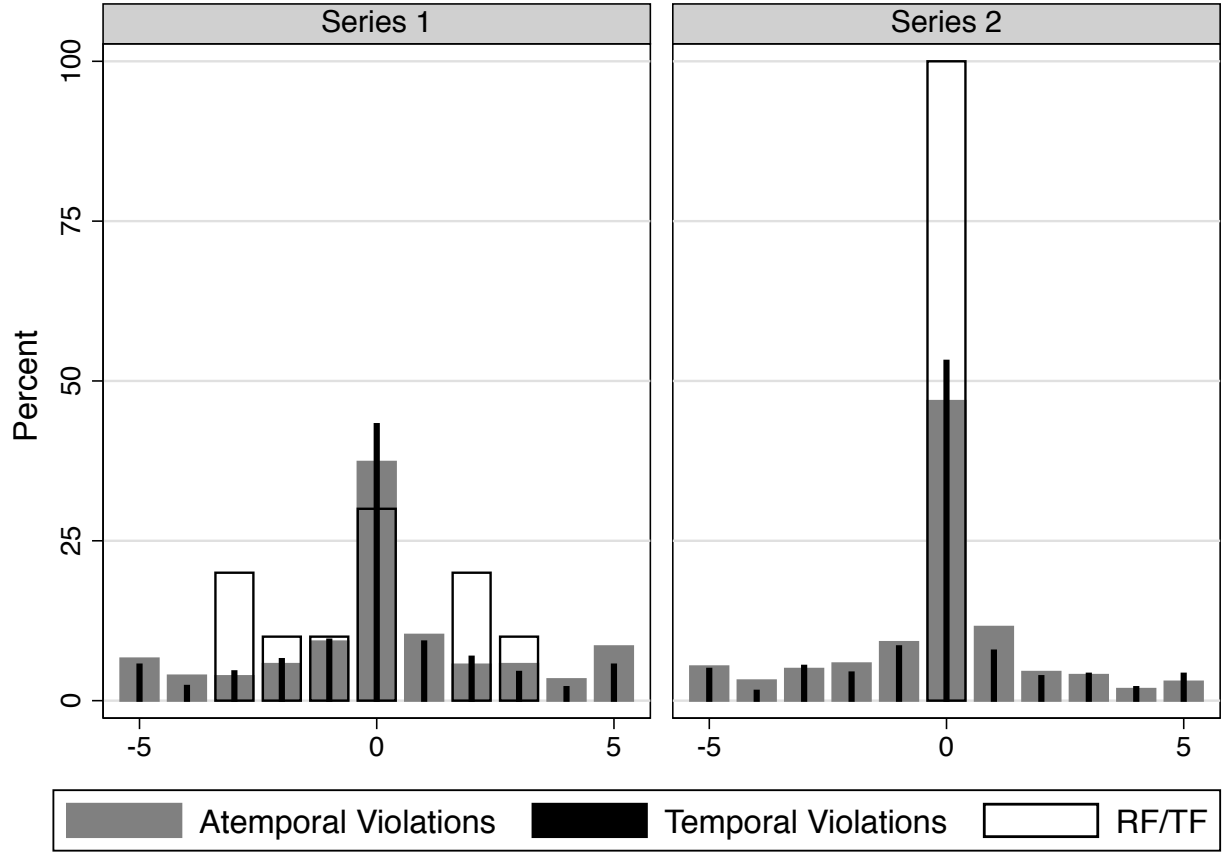


Figure A6: CRI Violations Without Time

*Notes:* Histograms of Atemporal Common Risk Invariance violations for Series 1 (Panel A) and Series 2 (Panel B). 10 comparisons per series  $\times$  105 subjects = 1050 observations in each panel. Actual data as grey bars. Time-First Prediction/Risk-First predictions as white bars. Temporal Common Future Risk Invariance violations as black lines for comparison. All comparisons calculated difference in number of Option B choices between task  $j$  and task  $k$  with aggregate prediction for  $j < k$ . Negative numbers correspond to increased risk aversion, positive numbers to increased risk tolerance.

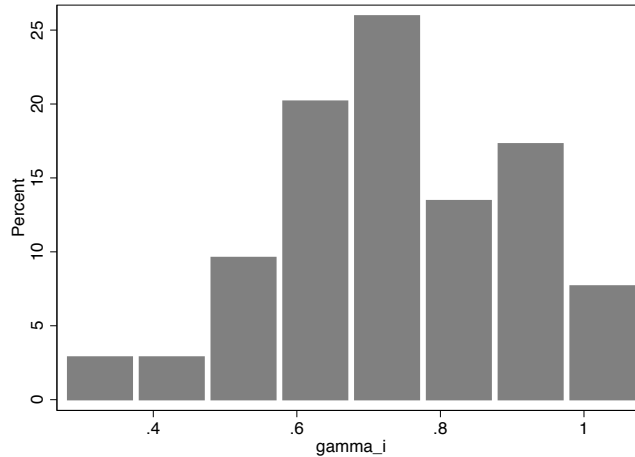


Figure A7: Individual Probability Weighting Estimates

*Notes:* Distribution of probability weighting parameters by subject. Parameters estimated from non-linear regression from the seven prospect theory evaluation tasks for every subject.

## A.4 Experimental Instructions

### Experimental Instructions

Please read carefully

Hello and Welcome.

#### **ELIGIBILITY FOR THIS STUDY:**

To be eligible for this study, you **must** have a mailing address. This address must be a valid way for you to receive mail from now through the end of Spring Quarter.

You must also be willing to receive your payment for this study by check. Checks will be written by Professor Charles Sprenger, Associate Professor of Economics. The checks will be drawn on the CHASE bank on campus. You may deposit or cash your checks wherever you like. If you wish, you can cash your checks for free at CHASE any weekday from 9:00 am to 6:00 pm or Saturdays from 10:00 am to 2:00pm at the Price Center with valid

identification (drivers license, passport, etc.).

Checks will be delivered to you via the mail. The value and timing of payments you will receive will be determined by your decisions in this study, and by chance. The soonest you could be mailed a payment is in a week. The latest you could be mailed a payment is four weeks from today. If you do not meet all of these criteria, please inform us of this now.



## Earning Money:

You will receive a \$10 participation payment. This payment will be split in two and mailed to you at two different dates. You will be mailed \$5 in one week and another \$5 in four weeks. These payments are independent of any other payments you receive from this experiment.

In this study, you will complete 29 different Tasks. Each task asks you to make a series of decisions between two options, OPTION A and OPTION B. The different options provide different values of payments sooner, mailed in one week, and later, mailed in four weeks. For each decision, all you have to do is decide whether you prefer OPTION A or OPTION B.

It is important to note that the payments in this study may involve chance. There may be a chance that your sooner payment, your later payment or both will not be sent at all. For each decision, you will be fully informed of the chance involved for the sooner and later payments. You will also always be fully informed about how each chance will be determined for each Task and each Decision. Whether or not your payments will be sent will be determined at the END of the experiment today.

Once all decisions have been made, we will randomly select one task and one decision as the **decision-that-counts**. This will be done in two stages. First, we will pick a number from 1 to 29 at random to determine which is the task-that-counts. Then we will pick a second number at random from 1 to 30 to determine which decision in the task-that-counts will be the decision-that-counts. We will use your choice in the decision-that-counts to determine your actual earnings.

*Note, since all decisions are equally likely to be chosen, you should make each decision as if it will be the decision-that-counts.*

**IMPORTANT:** All payments will be sent to your mailing address. On the scheduled day, a check will be placed for delivery with the United States Postal Service by Professor Sprenger and his assistants.

As a reminder to you, the day before you are scheduled to be mailed one of your payments, we will send you an e-mail notifying you that the payment will be sent. On your desk is a business card for Professor Sprenger with his contact information. Please keep this in a safe place. If one of your payments is not received you should immediately contact Professor Sprenger, and we will hand-deliver payment to you.

**Your Identity:** In order to receive payment, we will need to collect your name, mailing address and email address. This information will only be seen by Professor Sprenger and his assistants. As soon as your payments are registered today, the link between the choices you make and your payments will be destroyed. Your identity will not be a part of any subsequent data analysis. You have, instead, received an anonymous participant number on the card on your desk that we will use for analysis.

On your desk are two envelopes: one for the sooner payment and one for the later payment. Please take the time now to address them to yourself at your preferred mailing address. You will note that the participant number is written on both envelopes. Once your payments are registered today, we will erase this number on the envelope, destroying the link between your choices and your payments.

## How it Works:

In this study you will complete 29 Tasks. Each task asks you to make 30 decisions between two options. The first option will always be called OPTION A. The second option will always be called OPTION B. Each decision you make is a choice. For each decision, all you have to do is decide whether you prefer OPTION A or OPTION B.

Throughout the tasks, either OPTION A, OPTION B or both will involve chance. You will be fully informed of the chances involved for every decision. For this study, choices will be between various chances of a positive amount or **zero**. Options may have more than one chance of receiving a positive amount.

Once we know which is the decision-that-counts, and whether you prefer OPTION A or OPTION B, we will then determine the value of your payments. If you prefer OPTION A in the decision-that-counts, then OPTION A will be implemented. If you prefer OPTION B, then OPTION B will be implemented.

For example, consider the following two OPTIONS:

### EXAMPLE

Option A			or	Option B		
\$10 in a week Chance 1	\$20 in 4 weeks Chance 2			\$20 in a week Chance 1	\$20 in 4 weeks Chance 2	
1) 70 in 100 Chance	90 in 100 Chance	<input type="checkbox"/>	or	35 in 100 Chance	90 in 100 Chance	<input type="checkbox"/>
<i>if you prefer Option A, check the green box...</i>						
1) 70 in 100 Chance	90 in 100 Chance	<input checked="" type="checkbox"/>	or	35 in 100 Chance	90 in 100 Chance	<input type="checkbox"/>
<i>if you prefer Option B, check the blue box...</i>						
1) 70 in 100 Chance	90 in 100 Chance	<input type="checkbox"/>	or	35 in 100 Chance	90 in 100 Chance	<input checked="" type="checkbox"/>

If this was chosen as the decision-that-counts, four pairs of ten-sided dice would be rolled. One pair of dice would be rolled for **Chance 1** of **Option A**, one pair for **Chance 2** of **Option A**, one pair for **Chance 1** of **Option B**, and one pair for **Chance 2** of **Option B**. The rolls of these dice will determine the values of payments mailed in one week and in four weeks for those preferring **Option A** and those preferring **Option B**.

For each pair of dice, the first roll will correspond to the ten's digit, while the second roll will determine the one's digit. For example, rolling a 7 and then a 3 would mean rolling a 73. Rolling 0-0 will count as 100. If the chance of receiving a given amount is greater than or equal to the rolled number, then the corresponding amount will be mailed at the given date in addition to the \$5 participation payment. If the chance of receiving a given amount is lower than the rolled number, the corresponding amount would not be sent and only the \$5 participation payment would be mailed. For example, if the chance of receiving a given payment was 90 in 100, any dice roll from 0-1 to 9-0 will result in the payment being sent, while any dice roll from 9-1 to 0-0 will result in the payment not being sent.

In the example above, if you preferred **OPTION B** we would roll two pairs of dice for **Chance 1** and **Chance 2**. If the pair for **Chance 1** read 4-7 and the pair for **Chance 2** read 6-5, what payments would be mailed (don't forget your participation payments)?

For **Chance 1** you have a 35 in 100 chance of receiving \$20 in one week. Because 47 is larger than 35, the \$20 would not be received and you would be mailed  $\$0 + \$5$  participation payment = \$5 in one week. For **Chance 2** you have a 90 in 100 chance of receiving \$20 in four weeks. Because 65 is smaller than or equal to 90, the \$20 would be received and you would be mailed \$20

+ \$5 participation payment = \$25 in four weeks.

In the example above, if you preferred **OPTION A** we would roll two pairs of dice for **Chance 1** and **Chance 2**. If the pair for **Chance 1** read 6-8 and the pair for **Chance 2** read 9-8, what payments would be mailed (don't forget your participation payments)?

In one week:

In four weeks:

## Things to Remember:

- You must have a valid mailing address.
- You will receive your payment for this study by mail.
- You will receive a \$10 participation payment. This payment will be split in two. \$5 will be mailed in one week. \$5 will be mailed in four weeks. These payments are independent of the choices you make today.
- You will complete 29 tasks.
- Each task asks you to complete a series of decisions between OPTION A and OPTION B. The different options provide different values of payments sooner, mailed in one week, and later, mailed in four weeks. All you have to do is state whether you prefer OPTION A or OPTION B.
- In each task OPTION A, OPTION B or both may involve chance. You will be fully informed of the chances for each option.
- Options may have more than once chance of receiving a positive amount.
- Once all of your decisions have been made, we will choose one task and one decision as the **decision-that-counts** and will implement your preferred option.
- Every decision is equally likely to be the decision-that-counts. So, it is in your interest to treat each decision as if it could be the one that determines your payments.
- For the decision-that-counts, your payments may be determined by chance through the roll of four pairs of ten-sided dice. If a given chance payment will not be received, you will only receive the \$5 participation payment on the given date.